

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

IDENTIFICATION ET CARACTÉRISATION DE LA FAMILLE
DE FACTEURS DE TRANSCRIPTION AP2/EREBP CHEZ LE BLÉ

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN BIOLOGIE

PAR
BARBARA BOUCHO

AOÛT 2006

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 -Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire. »

REMERCIEMENTS

Je tiens d'abord à remercier mon directeur de recherche, le Dr Fathey Sarhan de m'avoir accueilli dans son laboratoire. J'ai pu bénéficier d'un soutien académique et d'un encadrement qui m'ont permis de croître tant au niveau professionnel que personnel.

Une reconnaissance particulière au Dr Jean Danyluk pour ses précieux conseils. Merci au Dr Normand Chevrier, au Dr Mario Houde, au Dr François Ouellet et tout le personnel du département des sciences biologiques (Université du Québec à Montréal).

Merci à tous mes collègues de laboratoire et ami(e)s qui m'ont aidés tout au long de ma maîtrise.

Un gros merci à mes parents pour leur amour et leur soutien.

A vous tous, merci

TABLE DES MATIÈRES

LISTE DES FIGURES	v
LISTE DES ABRÉVIATIONS.....	vii
RÉSUMÉ.....	viii

INTRODUCTION ET REVUE DE LITTÉRATURE.....	1
1. Introduction.....	1
2. Dommages causés par le gel.....	2
3. L'acclimatation au froid.....	4
3.1 Modifications de la membrane plasmique.....	4
3.2 Modification au niveau moléculaire.....	6
4. Les facteurs de transcription AP2/EREBP.....	9
4.1 La sous-famille APETALA 2 (AP2).....	11
4.2 La sous-famille RAV.....	12
4.3 La sous-famille ERF.....	13
4.4 La sous-famille CBF/DREB.....	14
4.5 Autres : dernière sous-famille non définie.....	17
5. Problématique.....	18
 CHAPITRE I : ARTICLE.....	 20
Title.....	21
Summary.....	22
Introduction.....	23
Experimental procedures.....	26
Results.....	31

Discussion.....	36
References.....	43
Legends to figures.....	49
Figures.....	53

CHAPITRE II : EXPRESSION ET PURIFICATION DE LA PROTEINE

<i>Ta</i> CBF1.....	64
Conclusion.....	72
Annexes.....	74
Références.....	91

LISTE DES FIGURES

Tableau 1: wheat genes containing the AP2domain.....	53
Figure 1: Phylogenetic analyses of wheat <i>CBF</i> genes.	
Phylogenetic analyses of 35 <i>CBF</i> genes from hexaploid wheat.....	55
Figure 2: Evolution of the <i>CBF</i> subgroup in plants.....	56
Figure 3: B. Structure and sequence alignment of the 15 <i>TaCBFs</i> with other <i>CBFs</i> members.....	57
Figure 4: Mapping of <i>TaCBF</i>	
A. Schematic representation of the deletion lines used to map <i>TaCBF</i> genes on the wheat chromosomes 2.....	59
B. Schematic representation of the deletion lines used to map <i>TaCBF</i> genes on the wheat chromosomes 5.....	59
Figure 5: Accumulation of <i>TaCBF</i> 2, 4, 9, 12, 13 transcripts during cold acclimation in winter wheat cv Norstar and spring wheat cv Quantum.	
A. Northern blot.....	60
B. Quantification of <i>TaCBF</i> 4 from group 1.....	60
Figure 6: Accumulation of <i>TaCBF</i> 1, 5, 7, 8 transcripts during cold acclimation in winter wheat cv Norstar and spring wheat cv Quantum.	
A. Northern blot.....	61
B. Quantification of <i>TaCBF</i> 1 from group 2.....	61

Figure 7: Accumulation of *TaCBF11* transcripts during cold

acclimation in winter wheat cv Norstar and spring wheat cv Quantum.

- A. Northern blot62
- B. Quantification of *TaCBF11* from group 3.....62

Figure 8: Accumulation of *TaCBF3*, 9, 11 transcripts during coldacclimation in winter wheat cv Norstar and spring wheat cv
Manitou.....63

LISTE DES ABRÉVIATIONS

ADN	acide désoxyribonucléique
ADNc	acide désoxyribonucléique complémentaire
ARN	acide ribonucléique
ARNm	acide ribonucléique messenger
bp	base pair
CBF	CRT-binding factor
CRT	C-repeat
cv	cultivar
DRE	dehydration-responsive element
DREB	DRE-binding protein
LT	low temperature
PCR	polymerase chain reaction
QTL	quantitative trait locus
SDS	sodium dodecyl sulfate
SSC	sodium saline citrate
v/v	volume/volume
w/v	weight/volume

RÉSUMÉ

Pendant leur cycle de vie, les plantes sont exposées aux différents stress abiotiques tels que les basses températures, la sécheresse et les grandes concentrations de sel dans le sol. Ces facteurs sont connus pour changer le métabolisme des plantes et l'expression des gènes. L'étude de certains de ces gènes régulés par le froid ainsi que les facteurs de transcription avec lesquels ils interagissent ont permis de révéler que le sous-groupe de gènes CBF, appartenant à la famille de facteurs de transcription AP2, joue un rôle central dans la régulation de l'expression des gènes au froid. Bien que plusieurs études aient permis de montrer le potentiel de ces facteurs dans l'amélioration de la tolérance aux stress, nos connaissances restent minimes sur cette famille de gène. Avec le but à long terme de comprendre les fonctions des différents membres de cette famille dans le blé, nous avons initié cet étude afin d'identifier tous les gènes CBF de blé (TaCBF) et de caractériser leurs profils d'expression sous divers stress abiotiques. En utilisant des criblages de banques de ADNc de blé, des recherches dans des banques de données et des amplifications par PCR, 35 ADNc de CBF ont été identifiés chez le blé hexaploïde. Les profils d'expression (northern et RT-PCR) ont révélé que la plupart des gènes TaCBF sont induits au froid et sont exprimés plus fortement chez le génotype d'hiver comparativement aux génotypes de printemps. L'analyse phylogénétique révèle que le blé contient au moins 15 gènes CBF différents comparé aux 10 gènes présents chez le riz et 6 chez *Arabidopsis*. Le plus grand nombre de gènes contenus dans le blé peut être responsable de sa plus grande tolérance au gel. La compréhension de la fonction et de l'évolution de cette famille multigénique chez le blé aidera à fournir les bases pour des stratégies d'ingénierie qui mèneront à des plantes ayant une plus grande tolérance aux stress.

Mots clés : blé hexaploïde, acclimatation au froid, facteurs de transcription AP2, CBF.

INTRODUCTION ET REVUE DE LITTÉRATURE

1. Introduction

Au cours de leur évolution, de nombreuses plantes ont développé des mécanismes qui leur permettent de s'adapter à différents stress environnementaux, notamment des températures élevées, des températures basses (froid, gel), des excès d'eau (anoxie), la sécheresse, la salinité, des radiations excessives ou des stress chimiques et biotiques. Le froid, par exemple, est un facteur limitant pour l'extension des cultures dans les régions nordiques, et le gel est un facteur pouvant avoir un impact considérable en agriculture. Les plantes peuvent répondre aux stress de différentes façons : elles peuvent s'adapter aux conditions de stress et développer une tolérance ou elles peuvent subir des dommages et, finalement, périr. En fait, des modifications spécifiques du métabolisme permettraient à certaines plantes d'éviter ou de tolérer les impacts des stress (Gilmour *et al.*, 1998). Ces modifications comprennent des changements biochimiques, moléculaires, biophysiques ainsi que physiologiques qui semblent capitales pour développer la résistance aux stress (Thomashow, 1990). Par exemple, du seigle qui a été soumis à un régime d'acclimatation à de basses températures (au-dessus de 0°C) peut par la suite survivre jusqu'à -30 °C, alors que la même plante mourrait à -5°C si elle était préalablement exposée seulement à des températures de croissance normale (20°C).

La capacité des plantes à subir le changement des saisons est déterminante pour obtenir un bon rendement lors des récoltes. Afin de comprendre comment les plantes sont capables de s'acclimater à de basses températures, il est important de connaître d'abord la nature des dommages causés par le gel.

2. Dommages causés par le gel

De nombreuses études ont indiqué que le système membranaire de la cellule chez la plante était le premier endroit où les dommages causés par le gel sont détectés. La glace se forme dans l'espace intercellulaire dès que la température baisse en dessous de 0 °C. Cette formation est due à la différence de point de congélation qui existe entre le fluide extracellulaire (plus élevé) et le fluide intracellulaire (Thomashow, 1999). Il se crée alors un mouvement d'eau liquide de l'intérieur de la cellule vers l'espace intercellulaire car le potentiel chimique de la glace est plus faible que celui de l'eau liquide à la même température (Thomashow, 1990). Il en résulte un déséquilibre osmotique.

La déshydratation due au gel peut causer de nombreux dommages à la membrane. Parmi les plus importants, il y a les transitions de la phase lamellaire à la phase hexagonale II (lamellar-to-hexagonal-II phase transitions) et les lyses induites par l'accroissement cellulaire (expansion-induced-lysis).

Les transitions de la phase lamellaire à la phase hexagonale II augmentent la pression osmotique à l'intérieur des différentes couches de la membrane. Ce qui engendre une grande pression osmotique entre les différentes couches de cellules autour de la membrane plasmique. Cette pression dépasse la force qui, normalement, sépare les différentes couches de cellules et il y a donc plus d'espace entre les différentes couches (Uemura et Steponkus, 1994).

Les lyses induites par l'accroissement cellulaire sont une conséquence de la grande fluctuation osmotique qui survient lors du cycle de congélation/décongélation. Quand le potentiel osmotique diminue lors de la formation de gel, il y a apparition de vésicules endocytotiques et il en résulte un rétrécissement de la surface de la

membrane plasmique. Ces diminutions de surface sont irréversibles. Il en résulte une augmentation du potentiel osmotique qui prend place durant la période de décongélation, la membrane cellulaire se déchire avant d'avoir retrouvé son volume initial (Uemura et Steponkus, 1994).

Même si les dommages sont généralement causés par une déshydratation cellulaire, plusieurs autres facteurs peuvent contribuer à ces dommages. De nombreuses études ont montré que le gel entraînait la formation de molécules d'oxygène réactives comme le peroxyde d'hydrogène qui cause des lésions à la membrane plasmique (Tao *et al.*, 1998). De plus, il est connu que les basses températures induisent la dénaturation des protéines, ce qui pourrait, à long terme, entraîner des dommages cellulaires. Cela aura pour conséquence un ralentissement de croissance pouvant mener à la mort de la plante (Guy *et al.*, 1998). Des changements sur le plan de la fluidité de la membrane sont aussi un des effets directs du froid sur la plante ce qui continue à renforcer l'hypothèse que la membrane aurait le rôle de percevoir le signal des basses températures (Levitt, 1980).

Compte tenu que les premiers dommages observés au niveau cellulaire sont dus à une importante déshydratation, il serait logique de penser que la tolérance au froid soit associée à une tolérance à la déshydratation. Une des fonctions majeures de l'acclimatation au froid serait donc la stabilisation de la membrane contre les dommages de déshydratation causés par le gel.

3. L'acclimatation au froid

Chez les plantes tolérantes au gel, les nombreuses modifications qui prennent place lors de l'exposition aux basses températures et qui conduisent à une tolérance accrue au gel sont collectivement appelées acclimatation au froid. C'est un processus qui permet à la plante de minimiser les dommages et de survivre aux températures extrêmes. Ce processus est déclenché par une période de croissance à de basses températures au-dessus du point de congélation. La durée de cette période d'acclimatation au froid peut varier d'une espèce de plante à l'autre. Chez la pomme de terre, l'épinard et *Arabidopsis*, elle varie entre 10 et 15 jours (Gilmour *et al.*, 1988) et chez le blé et le seigle, elle est d'environ 50 jours (Fowler *et al.*, 1996).

3.1 Modifications de la membrane plasmique

La différence majeure qui existe entre les membranes de plantes acclimatées et non acclimatées est que le matériel membranaire reste intact au cours du cycle de gel et dégel chez les cellules tolérantes. Les cellules des plantes tolérantes sont capables de modifier leur paroi et la membrane plasmique pour protéger la plante des blessures causées par le gel. C'est pour cela que les dommages comme les transitions de la phase lamellaire à la phase hexagonale II (Hrr) et les lyses induites par l'accroissement cellulaire sont beaucoup moins nombreux. En effet, les protoplastes de plantes acclimatées ne forment pas de vésicule endocytotique qui mènerait inévitablement à une perte de surface. En revanche, il y a formation d'extrusions exocytotiques qui permettent à la membrane de retrouver sa surface initiale sans subir de déchirement lors de la période de dégel (Uemura *et al.*, 1995).

Plusieurs études ont montré que, lors d'un stress au froid, les teneurs en stérols et en phospholipides de la membrane plasmique augmentent. Par contre, les teneurs en stérols glucosides acétylés et cérébrosides (CER) diminuent. Ces modifications ont été observées chez *Arabidopsis*, le seigle et l'avoine lors de l'acclimatation au froid. Ces changements auraient un rôle dans la cryostabilité de la membrane plasmique lors du gel (Uemura *et al.*, 1995). L'accroissement en phosphatidylcholine, qui est le lipide le plus hydraté des espèces présentes dans la membrane plasmique, et la diminution en cérébrosides, qui sont les espèces les moins hydratées, auraient un effet direct sur la membrane. Tout d'abord, cela entraînerait l'augmentation des propriétés hydratantes de la membrane plasmique et par conséquent empêcherait l'accolement de la membrane plasmique avec les autres membranes pendant la déshydratation. Puis, il y aurait diminution des échanges lipide-lipide entre les deux couches de cellules (Webb *et al.*, 1993). Cela aurait un effet direct sur la courbure de la monocouche lipidique et donc une diminution de la formation de la phase Hrr (Uemura *et al.*, 1994). Ces modifications sont associées à une réduction de la quantité d'eau présente à l'intérieur des cellules, une réduction du volume des cellules, une augmentation des solutés intracellulaires et une augmentation de la rigidité de la plante (Guy, 1990). Elles ont donc un effet direct sur la fluidité de la membrane. Des accumulations de saccharose, raffinose, sorbitol et fructose ont aussi été observées durant l'acclimatation au froid chez les plantes. Selon Crowe and Crowe (1992), les sucres auraient la capacité de maintenir la stabilité structurale des protéines et des membranes déshydratées dans les mêmes conditions physiologiques que celles observées dans l'eau. Des sucres seraient aussi capables de former une phase vitrifiée en absence d'eau. Cette phase serait capable de ralentir les réactions chimiques et stabiliser le métabolisme de la plante lors d'un stress (Crowe *et al.*, 1997).

3.2 Modifications au niveau moléculaire

Les études génétiques classiques ont montré que l'acclimatation au froid chez les plantes était un caractère quantitatif impliquant l'action de plusieurs gènes ayant des effets additifs (Thomashow, 1990). Au cours des dernières années, trois approches importantes ont été utilisées pour déterminer la nature de gènes qui ont un rôle dans l'acclimatation au froid : l'isolement et la caractérisation de gènes régulés pendant l'acclimatation; l'isolement et la caractérisation de mutants associées à la tolérance au froid, et la détermination des « quantitative trait locus » (QTL) ayant un rôle dans la tolérance (Thomashow, 1990).

Selon l'hypothèse de nombreux chercheurs, le mécanisme par lequel la plante commencerait à s'acclimater, serait le suivant: 1- le froid est détecté par la membrane plasmique ou le chloroplaste, 2- un signal sera ensuite envoyé grâce à une cascade de kinases et de phosphatases, 3- cette cascade va causer l'activation de facteurs de transcription spécifiques qui, à leur tour, vont activer la transcription de nombreux gènes COR « cold-regulated » en se liant au motif de liaison DRE/CRT présent dans leur promoteur. L'expression de ces gènes serait alors responsable du développement de la tolérance au froid (Breton *et al.*, 2000).

Des études ont permis de classer les protéines identifiées en quatre catégories différentes selon leur fonction (Breton *et al.*, 2000).

- 1- Les protéines structurales permettant de protéger les tissus végétaux contre différents stress (protéines de l'embryogenèse (LEA), protéines antigel, osmotine, protéines liants ARNm).
- 2- Les enzymes qui ont un rôle dans la biosynthèse de différents osmoprotecteurs, dans la désaturation de lipides et la réponse antioxydative.

- 3- Les protéines de régulation qui ont un rôle dans l'expression des gènes et dans la transduction des signaux (sous-famille DREB, sous-famille ERF, zinc finger, BZIP, MADS, kinases, phosphatases).
- 4- Les protéines qui ont des fonctions inconnues jusqu'ici.

Le premier groupe de protéines inclut une grande famille de protéines qui sont les LEA. Ces polypeptides sont synthétisés à la fin de l'embryogenèse, ce qui leur confère leur nom, leur nature est hydrophile et plusieurs contiennent une région capable de former une hélice alpha amphipathique (Thomashow, 1999). Ils sont aussi souvent associés à un stress hydrique, ce qui a permis d'émettre l'hypothèse selon laquelle ces gènes auraient un rôle dans la tolérance au froid en protégeant les membranes de la cellule contre les dommages causés par la déshydratation. Par exemple, la protéine *WCS120* chez le blé (Houde *et al.*, 1992) et *HVA1* chez l'orge (Hong *et al.*, 1992) font partie des groupes II et III de la famille LEA. Sarhan *et al.* (1997) ont étudié l'expression du gène *WCS120* au cours de l'acclimatation au froid chez différents cultivars de blé. Les résultats ont montré que le niveau d'induction du gène à 4 °C chez les cultivars de printemps était inférieur à celui observé chez les cultivars d'hiver. De plus, le niveau d'expression du gène *WCS120* est corrélé avec le développement de la tolérance au froid. D'autres protéines homologues aux protéines LEA comme la protéine *COR15a* chez *Arabidopsis* (Lin et Thomashow, 1992) et *Cas15* chez la luzerne (Monroy *et al.*, 1993) sont aussi associées à la tolérance au froid. Les gènes *COR* (Cold-Regulated) comptent dans leur région promotrice plusieurs copies d'éléments *cis* nommés DRE/CRT (Dehydration Response Element / C-Repeat binding factor) (Yamaguchi-Shinozaki et Shinozaki, 1997; Thomashow, 1999). Ces éléments sont liés par des facteurs de transcription lors d'une exposition à basse température. Ces facteurs ont été clonés chez *Arabidopsis*, le blé, le seigle, l'orge et la tomate (Jaglo *et al.*, 2001; Choi *et al.*, 2002). Il y a aussi les protéines antigels ou AFP (antifreeze proteins) qui s'accumulent dans l'apoplaste, où les premiers cristaux de glace se forment lors de l'acclimatation (Antikainen et Griffith,

1997). Ces protéines auraient un rôle dans l'inhibition de la formation des cristaux de glace entre les cellules de la plante.

Le deuxième groupe de protéines inclut des enzymes qui ont un rôle dans la réponse antioxydative, la biosynthèse de différents osmoprotecteurs et dans la désaturation de lipides. Les espèces réactives d'oxygène (ROS), dont la formation est accélérée dans les conditions de stress, doivent être rapidement neutralisés pour éviter des dommages oxydatifs. La durée de vie des ROS ou des radicaux libres dans l'environnement cellulaire est déterminée par le système antioxydatif, qui fournit la protection cruciale contre le dommage oxydatif. Une protection contre les dommages oxydatifs est aussi assurée par des métabolites (vitE...) et par de nombreuses enzymes. Une étude réalisée pour comparer les enzymes des tissus acclimatés au froid par rapport à celles des tissus non acclimatés a montré que l'activité de la peroxydase augmentait dans les tiges acclimatées (Guy, 1990). La superoxyde dismutase (O_2^-) est une enzyme qui catalyse la transformation de l'anion superoxyde en O_2 et H_2O_2 et réduit ainsi le niveau des espèces réactives d'oxygène dans la cellule (McKersie *et al.*, 1993). Des niveaux élevés de molécules enzymatiques comme l'ascorbate peroxydase, la glutathion réductase, la monodehydroascorbate réductase et la déhydroascorbate réductase s'accumulent chez le pin pendant l'acclimatation au froid (Tao *et al.*, 1998).

Le troisième groupe de protéines inclut de nombreuses protéines qui ont un grand rôle dans la régulation de l'expression des gènes et dans la transduction des signaux. Par exemple, la protéine *TaVRT-1* qui fait partie de la famille des MADS chez le blé (Danyluk *et al.*, 2003) a un rôle important dans la vernalisation et la protéine kinase *OsCPK2* chez le riz est détectée dans les membranes lors de l'acclimatation (Martin et Busconi, 2000). L'expression des protéines de la sous-famille DREB est aussi observée chez les blé lors de l'acclimatation ce qui suppose un rôle au niveau de l'acquisition de la tolérance au froid.

4. Les facteurs de transcription AP2/ EREBP

Le clonage des facteurs de transcription CBF1/DREB1A par les équipes de Thomashow (1999) et Shinozaki (2000) a été un point tournant important dans le domaine de la génétique moléculaire sur l'acclimatation au froid. L'expression de *CBF1* chez des plantes transgéniques d'*Arabidopsis* a été associée à une régulation positive de l'expression des gènes *COR*. Au cours de cette expérience, il a été démontré que ces plantes avaient acquis une augmentation importante de leur tolérance au froid.

Les facteurs de transcription végétaux sont donc des protéines de régulation qui activent ou inactivent les mécanismes plurigénétiques, qui ont pour effet de changer de nombreuses propriétés chez la plante. Les gènes de facteurs de transcription sont des « masterswitches » naturels des mécanismes qui contrôlent plusieurs caractères importants chez les plantes.

Les membres de la famille de facteurs de transcription AP2/ EREBP ont différents rôles dans le cycle de vie de la plante, comme le développement des organes ou l'identité cellulaire. L'expression de ces gènes est influencée par de nombreux stress et leurs rôles sont cruciaux pour la survie de la plante. L'importance de leur rôle fait de ces gènes des sujets de recherche importants pour la compréhension de la tolérance aux stress abiotiques comme la tolérance au gel. C'est pour cela que Sakuma *et al* (2002) ont entrepris l'identification de 145 protéines DREB/ERF chez *Arabidopsis*, qui ont été classifiées en 5 sous-familles : AP2 (APETALA2), RAV, DREB, ERF et « autres » (tableau 1). Cette classification a été basée sur des similitudes aux acides aminés des domaines de liaison AP2.

Tableau 1: Classification des gènes de la famille AP2 chez *Arabidopsis* (Sakuma et al., 2002).

Results of Database Search for <i>Arabidopsis</i> Genes Containing the ERF/AP2 Domain				
DNA-binding domain type	Subgroup	Total genes of <i>Arabidopsis</i>	Number of expressing genes ^a	Representative genes
Double ERF/AP2	APETALA2 subfamily	14	7	APETALA2, AINTEGUMENTA, RAP2.7
ERF/AP2 + B3	RAV subfamily	6	3	RAV1, RAV2
Single ERF/AP2	DREB subfamily			
	A-1	6	4	DREB1/CBF
	A-2	8	2	DREB2
	A-3	1	1	ABI4
	A-4	16	3	TINY
	A-5	16	9	RAP2.1, 9, 10
	A-6	9	5	RAP2.4
	Subtotal	56	24	
	ERF subfamily			
	B-1	15	8	AtERF3, 4, 7, 8, 9, 10
	B-2	5	3	AtEBP, RAP2.2
	B-3	18	11	AtERF1, 2, 5, 6, ERF1
	B-4	7	2	RAP2.6
	B-5	8	2	
	B-6	12	3	RAP2.11
	Subtotal	65	29	
	Others	4	2	
	Single ERF/AP2 subtotal	125	55	
Total		145	65	

^a Genes that were registered in the database as isolated genes or for which an EST sequence with an identical DNA sequence to the gene was found in the database.

Le domaine AP2 est un motif d'environ 70 acides aminés présents dans la séquence des protéines de cette famille. Il est composé de deux grands éléments : élément YRG et élément RAYD (Jofuku et al., 1994; Riechmann et Meyerowitz, 1998). Les homologies structurales de ce domaine ont été conservées dans les 70 acides aminés sous forme de trois feuillets bêta (figure 1). À ce jour, de nombreuses protéines de la famille AP2 ont été identifiées chez différentes espèces dont le maïs (Qin et al., 2004), la tomate (Zhang et al., 2004), le canola (Gao et al., 2002) et le riz (Lee et al., 2004).

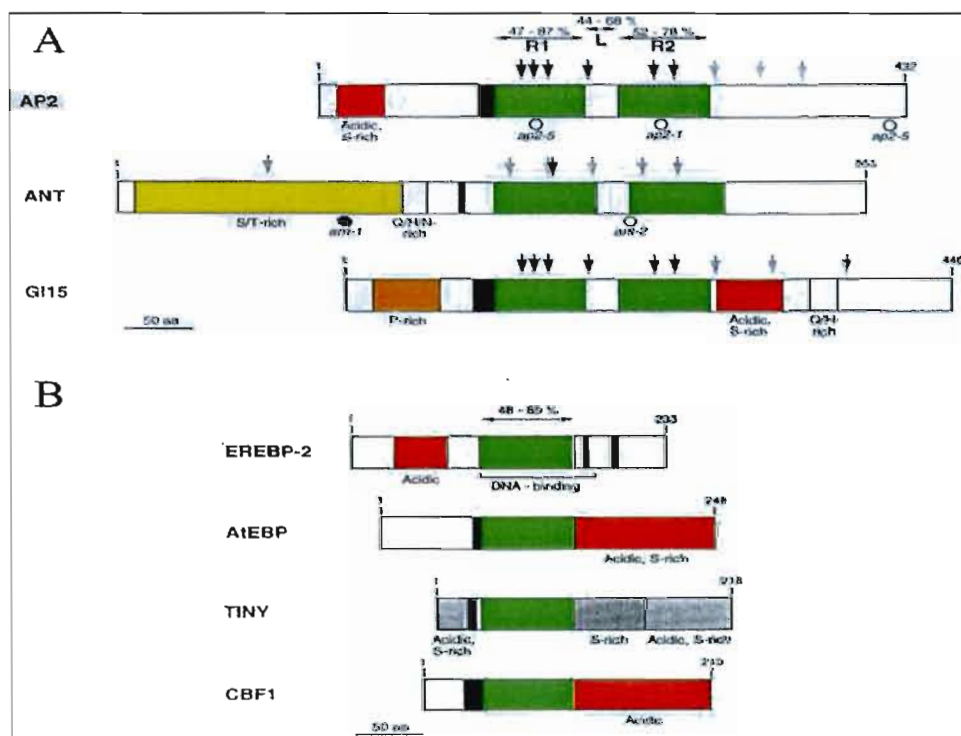


Figure 1 : Représentation des domaines caractéristiques à chaque sous-groupe de la famille AP2 (Riechmann et al., 2002). A= sous-famille DREB , B=sous-famille ERF.

4.1. La sous-famille APETALA 2 (AP2)

La sous-famille est composée de 14 membres chez *Arabidopsis* et de 20 membres chez le riz. Les gènes de cette famille jouent un rôle important dans le contrôle de la floraison et dans le développement de la graine. Ces gènes codent pour

des protéines contenant deux domaines de liaisons AP2 ainsi qu'un élément YRG qui consiste en 19 à 22 acides aminés (Okamuro *et al.*, 1997) (tableau 1).

Les études ont montré que l'ARNm des APETALA2 était différemment exprimé durant le développement de la plante dans les tissus floraux et végétatifs (Okamuro *et al.*, 1997). Les gènes APETALA2 ont un rôle important sur le plan de la régulation du développement car les mutants phénotypiques ont permis d'observer des effets sur le méristème, les organes de floraison, les ovules et l'enveloppe de la graine (Vahala *et al.*, 2001).

Le gène AINTEGUMENTA (ANT) a un rôle particulier dans la régulation de la division cellulaire lors du développement de l'ovule ainsi que dans la croissance de l'organe de floraison. L'absence d'ANT coïncide avec une diminution de taille à cause d'une baisse de la croissance cellulaire. L'utilisation de plantes de tabac transgéniques contenant 35S :: ANT montre une expression ectopique ainsi qu'un gain de la fonction ANT, grâce à une augmentation des organes embryogéniques et des tiges. (Mizukami et Fischer, 2000).

Le rôle de cette sous-famille est donc capital dans le développement de la plante au cours de son cycle de vie. Toutefois son rôle dans la réponse au stress reste encore inconnu.

4.2. La sous-famille RAV

Ces facteurs de transcription possèdent deux domaines de liaisons différents ERF/AP2 et B3. Le domaine B3 est conservé chez les homologues des facteurs AB13/VP1 (Sakuma *et al.*, 2002). Cette sous-famille est composée de six gènes chez *Arabidopsis* qui sont homologues aux gènes AB13/VP1 trouvés chez le maïs. Les gènes contenant un domaine VP1 jouent un rôle important dans le contrôle de gènes

régulés par l'acide abscissique. Chez le riz, 4 membres ont été identifiés à ce jour. La fonction et la cible des protéines de la sous-famille RAV restent inconnues. Deux protéines, RAV1 et RAV2 ont été le sujet de nombreuses recherches. La technique de gel de retardement a été utilisée pour des études de liaison. Cela a permis de voir que RAV1 se liait spécifiquement à deux motifs 5'-CAACA-3' et 5'-CACCTG-3' (Kagaya *et al.*, 1999). La liaison de deux domaines au lieu d'un seul montre une augmentation de l'affinité et de la spécificité de la liaison par rapport à la cible. Ces protéines sont retrouvées dans presque toutes les parties de la plante chez *Arabidopsis*.

4.3. La sous-famille ERF

Les gènes de cette sous-famille ont un profil d'expression régulé par de nombreux stress, notamment des maladies ou des infections dues à des pathogènes. Ces derniers régulent les facteurs de transcription grâce à des voies de signalisation comme l'éthylène, l'acide jasmonique ou l'acide salicylique (Gutterson et Zhang, 2004).

Les phytohormones jouent un rôle dans le développement et la réponse au stress en modulant l'expression de gènes. Par exemple, l'éthylène agit sur l'expression de gènes qui sont actifs lors d'attaque de pathogènes, de blessures, de basses températures ou de sécheresse (Wees *et al.*, 2000). Les protéines ERF contiennent un domaine de liaison AP2. Chez le riz, 83 membres ont été identifiés. Cette sous-famille contient 6 sous-groupes (B-1 à B-6) chez *Arabidopsis* regroupant 65 gènes qui ont un rôle dans les stress biotiques. La sur-expression de *Pti4* chez *Arabidopsis* induit l'expression de gènes contenant une boîte GCC et confère une résistance accrue aux attaques de pathogènes (Tournier *et al.*, 2003). De même, chez la tomate, le domaine ERF *Pti4/5/6*, présent sur le gène, se lie à une protéine kinase

reliée aux pathogènes (Tournier *et al.*, 2003). La liaison des facteurs de transcription ERF à la séquence cible AGCCGCC a été étudiée. Cet élément est trouvé dans de nombreux promoteurs de gènes induits par l'éthylène ou l'acide jasmonique et qui interviennent lors d'infection. Ces évidences permettent de conférer aux gènes de la sous-famille ERF un rôle important dans la protection de la plante contre les pathogènes et les dommages causés par de basses températures ou des sécheresses.

4.4. La sous-famille CBF/DREB

Les gènes de cette sous-famille sont induits par un ou plusieurs stress environnementaux et leurs produits jouent un rôle dans la régulation de la tolérance aux stress abiotiques chez les plantes. Ils sont composés d'un domaine de liaison AP2. Chez *Arabidopsis*, 56 membres ont été identifiés et chez le riz, 58. La sous-famille CBF/DREB est divisée en six sous-groupes en fonction de similarités dans leurs séquences (A-1 à A-6), les deux premiers groupes étant les plus étudiés.

Les protéines de liaison-DRE du sous-groupe A-1, CBF1, CBF2, CBF3, aussi connues sous le nom de DREB1B, DREB1C et DREB1A chez *Arabidopsis* sont des protéines de régulation qui ont un rôle important dans l'acclimatation au froid et dans l'expression des gènes *COR* (Medina *et al.*, 1999). Quand les plantes sont bien irriguées et poussent à une température normale (20 °C), aucun des gènes *CBF* ou *COR* n'est exprimé. Au cours de l'acclimatation, les gènes *CBF* sont induits rapidement dans la première demi-heure, suivis par les gènes *COR* (Stockinger *et al.*, 2001). La sur-expression des gènes *CBF1* ou *CBF3* dans des plantes transgéniques d'*Arabidopsis* résulte en l'expression des gènes *COR*, entraînant une augmentation de la tolérance au gel sans acclimatation au froid (Gilmour *et al.*, 2000). De plus, la sur-expression de *CBF3* produit l'accumulation de proline et de sucres solubles tels que le saccharose, le fructose et le glucose, de façon similaire à l'acclimatation (Gilmour *et al.*, 2000). En outre, les gènes induits par les CBF, comme le gène *COR15a* et les

gènes *LEA*, contribuent directement à l'augmentation de la tolérance. Ces résultats montrent une grande implication de ces gènes dans l'acclimatation au froid et la tolérance au gel.

Le deuxième sous-groupe de la famille CBF/DREB présente des caractéristiques importantes dans la tolérance à d'autres stress. Les protéines de liaison DRE du sous-groupe A-2, comme DREB2A, DREB2C et DREB2D chez *Arabidopsis*, sont des protéines de régulation qui ont un rôle important dans la tolérance à la déshydratation et à la salinité. Le gène de DREB2A est fortement induit dans les feuilles et les racines de plantes ayant subi un traitement de déshydratation et de salinité (Sakuma *et al.*, 2002).

Malgré la différence qui existe entre les stress de basses températures et la sécheresse, il a été démontré que de nombreux gènes communs sont induits par ces deux stress. En effet, le promoteur de plusieurs gènes induits au froid et à la déshydratation chez *Arabidopsis* contient un élément de régulation CRT (C-repeat)/DRE (dehydration-responsive element) (Baker *et al.*, 1994; Yamaguchi-Shinozaki et Shinozaki, 1994). L'élément CRT/DRE est présent dans le promoteur de gènes comme le gène *COR15a* (Artus *et al.*, 1996), le gène *COR6.6/KIN2* (Gong *et al.*, 2001) et le gène *COR47* (Puhakainen *et al.*, 2004) qui sont induits par le froid et la déshydratation. Donc, la sur-expression des protéines CBF/DREB1 chez *Arabidopsis* résulte en une augmentation de la tolérance au froid chez la plante non acclimatée et acclimatée (Jaglo-Ottosen *et al.*, 1998; Liu *et al.*, 1998; Kasuga *et al.*, 1999; Gilmour *et al.*, 2000) ainsi qu'une augmentation de la tolérance à la déshydratation causée par un déficit hydrique ou une exposition à un stress salin (Liu *et al.*, 1998; Kasuga *et al.*, 1999).

Une fois que l'identification de CBF a été faite, une nouvelle question est apparue. Comment les basses températures stimulent-elles l'expression de ces gènes?

S'appuyant sur le fait que les messagers de CBF commencent à s'accumuler dans les quinze premières minutes d'exposition à basse température, Gilmour *et al* (1998), émettent l'hypothèse que le détecteur de basses températures et « le transducteur de signal » sont présents à température pièce (25 °C) et donc ne suscitent pas le déclenchement de l'acclimatation. Il y aurait donc un facteur de transcription déjà présent à température pièce qui reconnaît les promoteurs CBF. Ce facteur ne serait pas une protéine CBF, car les promoteurs des gènes CBF ne contiennent pas l'élément de régulation CRT/DRE et la surexpression de CBF1 ne provoque pas d'accumulation de messagers CBF3. Gilmour *et al* (1998), ont suggéré que l'induction des gènes *COR* se ferait en deux étapes. L'induction de *CBF* impliquerait un activateur inconnu appelé ICE (inducer of CBF expression). ICE reconnaîtrait l'élément de régulation appelé Boîte ICE présent dans les promoteurs de chaque gène *CBF*. À température pièce, ICE est censé être dans un état inactif, soit parce qu'il est retenu dans le cytoplasme par une protéine de contrôle négative ou parce qu'il est dans une forme qui n'attache pas la Boîte ICE ou encore parce qu'il n'active pas la transcription des *CBF* efficacement. En revanche, , Gilmour *et al* (1998), présumant qu'il y aurait activation d'un signal qui modifierait ICE ou la protéine qui lui est associée suite à l'acclimatation; pour qu'à son tour ICE puisse déclencher l'expression du gène *CBF*. Il est toutefois possible qu'ICE ne puisse pas seulement réguler l'expression des gènes *CBF*, mais qu'il puisse aussi susciter l'expression d'autres gènes pouvant aussi avoir un rôle dans l'acclimatation au froid (Gilmour *et al.*, 1998).

De récentes études (Jaglo *et al.*, 2001) ont indiqué que le processus d'action des CBF était conservé chez *Brassica napus* et que certains composants de ce processus étaient présents chez le blé et le seigle qui s'acclimatent bien au froid, de même que chez la tomate qui, elle, ne s'acclimate pas. Les gènes de cette sous-famille sont d'une telle importance qu'ils auraient été conservés durant l'évolution, ce qui en

fait peut-être un des maillons les plus importants, sinon le plus important, dans le mécanisme d'acclimatation au froid chez la plante.

4.5. Autres: dernière sous-famille non définie

Les gènes contenant un ou deux domaines AP2 et qui n'ont aucune homologie avec les autres sous-familles sont classés dans cette sous-famille.

La famille AP2 s'est grandement étendue au cours de l'évolution grâce à de nombreuses duplications de gènes. Néanmoins, les séquences des domaines de liaisons ont été conservées et cela permet une meilleure identification des différents membres.

5. Problématique

Pour l'économie canadienne, la culture du blé est sur le plan économique la plus importante. Les cultivars de blé d'hiver, par exemple, sont connus pour produire de meilleurs rendements et des grains de meilleure qualité que les cultivars de printemps. Outre son importance économique, le blé possède une variabilité génétique qui en fait un excellent modèle pour l'étude de la tolérance aux stress. Dans les régions à climat tempéré, le froid constitue un facteur limitant pour le développement des cultures, ainsi que pour les gains économiques. Il est par conséquent important de développer des stratégies qui permettraient d'éviter les dommages causés par le froid.

Les outils génétiques d'aujourd'hui permettent de nouvelles stratégies dans l'identification de nouveaux gènes. Les informations autour de la structure et de la description de l'organisation du génome des plantes, le développement des cartes génétiques et physiques, et la cartographie fine des zones chromosomiques d'intérêt sont les outils de la biologie moderne qui ont pour but majeur l'identification de tous les gènes du génome. L'utilisation de ces techniques, bien qu'ayant présenté un taux de succès convenable, comporte ses limites. Le problème majeur du séquençage génomique par exemple, est la distinction peu évidente entre régions codantes et celles qui sont non-codantes et l'analyse fonctionnelle est dépendante du niveau d'expression des gènes ciblés. Il paraît donc nécessaire de perfectionner et de développer ces méthodes au maximum de leur potentiel pour favoriser les découvertes à venir.

L'analyse fonctionnelle et l'utilisation de gènes induits au cours du froid, sont utilisées afin de mieux comprendre les déterminants génétiques qui régulent la tolérance, la croissance et le développement des plantes. Chez *Arabidopsis* et d'autres plantes (blé, riz, ...) l'application de la génomique fonctionnelle a permis d'identifier

des gènes comme les facteurs de transcriptions AP2, qui s'avèrent de bons candidats dans les études de biologie moléculaire à cause de leur rôle dans l'expression de gènes lors de période de stress. Cette famille de facteurs de transcriptions est constituée de nombreuses sous-familles dont la sous-famille DREBP/CBF qui jouent un rôle important dans la tolérance au froid. Ils sont donc considérés comme des régulateurs clés dans le processus de la tolérance au froid. C'est pour cela que nos recherches se sont portées sur les membres de cette sous-famille.

Au cours de ce travail, le génotype Norstar du blé a été utilisé pour isoler et caractériser la famille de facteurs de transcription CBF. Nous avons donc essayer de déterminé (i) le nombre et la complexité des gènes CBF chez le blé, (ii) leurs contextes phylogénétiques comparées à d'autres céréales tolérantes et sensibles au froid, (iii) les caractéristiques d'expression de ces gènes en réponse aux stress abiotiques, et (iv) si les caractéristiques fonctionnelles des CBFs du blé se retrouvent chez les homologues du blé et chez *Arabidopsis*. Pour cela, de nombreuses techniques ont été utilisées : le criblage de banques d'ADNc ainsi que la collection FGAS (82000 EST), le clonage par PCR, la cartographie, des études phylogéniques et des profils d'expression obtenus par northern et Real Time PCR. Les recherches effectuées ont donc permis d'identifier 37 cDNAs de membres de la famille AP2/EREBP régulés par le froid chez le blé. Nos résultats ont permit d'améliorer nos connaissances sur la réponse au froid des monocots tolérants et ont démontré l'importance de prolonger des recherches sur des plantes modèles de non-récolte. Cela permettra par la suite d'appliqué les nouvelles connaissances aux récoltes qui sont économiquement importantes.

CHAPITRE I : ARTICLE

Contribution

Les travaux réalisés durant ma maîtrise ont été présentés sous forme d'un article scientifique. Pour les travaux associés à cet article, j'ai participé à plusieurs étapes du projet.

J'ai particulièrement été impliquée dans les manipulations concernant l'identification et la caractérisation moléculaire des gènes *TaCBFs*, à savoir le criblage de banques, les amplifications par PCR, et les profils d'expression (northern blots) présentés dans cet article. J'ai aussi participé à la rédaction du manuscrit, surtout les parties résultats et discussion.

Title

**WHEAT CBF GENE FAMILY: EXPRESSION PROFILING, MAPPING AND
EVOLUTIONARY ANALYSES**

M. Badawi, B. Boucho, J. Danyluk and F. Sarhan.

Département des Sciences biologiques, Université du Québec à Montréal, Case
Postale 8888, Succursale Centre-ville, Montréal, Québec, Canada H3C 3P8

Summary

During their life cycle, plants are exposed to abiotic stresses such as extreme temperature, drought, high light and high salt concentration in the soil. These stresses are known to alter plant metabolism and gene expression. The study of cold-regulated genes and transcription factors interacting with their promoters has revealed that the CBF (CRT-binding factor) subgroup (belonging to the AP2 transcription factor family) plays a central role in their regulation. Although several studies have shown the potential of these factors in improving stress tolerance in model plants, little is known about this family in the economically important species wheat. To understand the functions of the different members of this family in wheat, we set out to identify all wheat CBF (*TaCBF*) genes and characterize their expression profiles under various abiotic stresses. Using cDNA library screening, clones in the FGAS collection (Functional Genomics of Abiotic Stress, a Genome Canada project) and PCR amplification, we identified 37 genes from hexaploid wheat. Expression profiling using northern analysis and real time PCR revealed that most *TaCBF* genes are cold inducible and expressed to greater levels in winter habit genotypes compared to the spring ones. Phylogenetic analyses reveal that wheat contains at least 15 different genes compared to 10 genes present in rice and 6 in *Arabidopsis*. The amplification in gene content in wheat may, in part, be responsible for its superior freezing tolerance. Understanding the function and evolution of this multigene family in wheat will help to provide new strategies to create stress tolerant crops.

Introduction

Abiotic stresses like drought, high salinity and low/high temperature are environmental conditions that cause adverse effects on the growth of plants and the productivity of crops. Plants respond and adapt to these stresses at the molecular and cellular levels as well as at the physiological and biochemical levels. The expression of many genes is induced by these stresses in different plants species (Shinozaki and Yamaguchi-Shinozaki, 2000). The ability to tolerate cold temperatures varies greatly among plants (Sakai and Larcher, 1987). Plants from temperate regions, such as wheat, canola and *Arabidopsis*, are able to tolerate both chilling and freezing temperatures. In contrast, many species from tropical regions, such as tomato, maize and rice, are unable to tolerate freezing and suffer chilling injury when exposed to temperatures in the range of 0 to 12°C. Cold acclimation involves the activation of multiple mechanisms that contribute to an increase in freezing tolerance including the accumulation of low-molecular weight cryoprotective molecules, the synthesis of cryoprotective proteins and alterations in membrane lipid composition (Gilmour *et al.*, 2000). It has recently been established that an important component of cold acclimation in *Arabidopsis* is the CBF cold response pathway (Thomashow, 2001). Within 15 min of exposing plants to low temperature, a small family of transcriptional activators known as either CBF1, CBF2, and CBF3 (Stockinger *et al.*, 1997) or DREB1b, DREB1c, and DREB1a, are induced (Liu *et al.*, 1998). This induction is quickly followed by the production of CBFs downstream genes. The immediate target genes of the CBF/DREB1 proteins have CRT (C-repeat)/DRE (dehydration responsive element) elements in their promoters, the DNA regulatory sequence to which the CBF/DREB1 transcriptional activators bind (Yamaguchi-Shinozaki and Shinozaki, 1994). Overexpression of the CBF regulon in *Arabidopsis* results in an increase in freezing tolerance due, in part, to the induction of genes encoding cryoprotective proteins and enzymes involved in the synthesis of low-molecular weight cryoprotectants (Gilmour *et al.*, 1998). It also results in an increase

in tolerance to drought and high salinity (Kasuga *et al.*, 1999) and appears to have a role in acclimation to chilling temperatures (Gong *et al.*, 2002). The CBF/DREB1 proteins belong to the AP2/ERF family of DNA-binding proteins. This protein family, which comprises more than 144 members in *Arabidopsis*, has in common the AP2/ERF DNA-binding motif (Riechmann and Meyerowitz, 1998). However, CBF1, 2 and 3, and three other CBF/DREB1 proteins (CBF4/DREB1d, DREB1e and DREB1f) are distinguished from the rest of the AP2/ERF protein by a conserved set of amino acid sequences adjacent to the AP2/ERF domain. These sequences, PKK/RPAGRxKFXETRHP and DSAWR, have been designated as the CBF 'signature sequences' (Jaglo *et al.*, 2001) and their function remains to be determined. CBF-like proteins containing the signature sequences are present in a wide range of plants that cold acclimate (Jaglo *et al.*, 2001) as well as in plants that do not cold acclimate (Jaglo *et al.*, 2001). Moreover, genes encoding apparent CBF orthologs are induced rapidly in response to low temperature in *B. napus* (Jaglo *et al.*, 2001), barley (Choi *et al.*, 2002), tomato (Jaglo *et al.*, 2001) and rice (Dubouzet *et al.*, 2003). Thus, components of the CBF cold response pathway are highly conserved in plants. Gene duplications are generally considered as a mechanism for increasing diversity and function. After duplication, one gene copy may maintain the original function, whereas the other can be mutated and, thereby, could acquire a novel function (Zhang *et al.*, 2003). One example of potential adaptive divergence among regionally duplicated genes is the recently described CBF/DREB1 tandem array in *A. thaliana* (Gilmour *et al.*, 1998; Haake *et al.*, 2002; Medina *et al.*, 1999). *Cis*-acting and *trans*-acting elements, known collectively as CBF/DREB1, exist as three tandemly arrayed genes (*CBF1*, -2 and -3; the *CBF1-3* tandem array) occurring within an 8.7 kb region on chromosome IV (Gilmour *et al.*, 1998; Haake *et al.*, 2002; Medina *et al.*, 1999). The expression of these duplicated transcriptional activation genes are induced by low temperature and the products of these genes subsequently activate many other low temperature-responsive genes (Seki *et al.*, 2001). A copy of the *CBF/DREB1* genes, (*CBF4*) has been reported to be expressed under drought stress, but not cold

stress, while gene expression of the *CBF1-3* tandem array is activated by low temperatures. Further analysis of *CBF4* indicates that the divergence of this gene from the *CBF1-3* tandem array is not due to acquiring new function, but to the diversification of promoter regulatory elements responding to different, but closely related, environmental signals (Haake et al., 2002). *CBF4* is located on a different chromosome (chromosome V) and phylogenetic analysis indicates that an ancestral gene duplication, probably segmental duplication, preceded the split between the *CBF1-3* tandem array and *CBF4* (Haake et al., 2002). The divergence between the *CBF1-3* array and *CBF4* is coincident with divergent functional trajectories between these two clades.

To better understand the involvement of the CBF-pathway in wheat during cold stress, we initiated a study to identify the CBF gene family in hexaploid wheat. Here, we describe the general feature and molecular evolution of wheat CBF genes (*TaCBFs*) encoding proteins that are highly similar to the *Arabidopsis* CBF/DREB1 family.

Experimental procedures

Plant materials, growth conditions, and RNA purification

In this study we used spring wheat genotype (*Triticum aestivum* L. cv Quantum) and winter wheat genotype (*T. aestivum* L. cv Norstar). Wheat plants were germinated in moist sterilized vermiculite for 7 days under artificial light ($200 \mu\text{mol m}^{-2} \text{sec}^{-1}$) at 20°C/20°C (day/night). Control plants were maintained under the same conditions while cold acclimation and other stress treatments were performed as previously described (Frenette Charron *et al.*, 2002). Total RNA was extracted and subjected to another round of purification using RNeasy PlantMiniKit (QIAGEN, Hilden, Germany) to further ensure the purity of the RNA samples. A total of 15 plants were harvested on dry ice at the desired time point and immediately frozen at -70°C. After that total RNA was isolated using the RNeasy PlantMiniKit. Purified RNA (2.8 μg) was reverse transcribed in a 20 μl reaction volume using the Superscript II First Strand cDNA Synthesis System for RT-PCR (Invitrogen, Carlsbad, CA). Parallel reactions for each RNA sample were run in the absence of Superscript II (no RT control) to assess for any genomic DNA contamination. The reaction was terminated by heat inactivation at 70°C for 15 min. Subsequently, the cDNAs products were treated with 2 units of RNase H for 20 min at 37°C. cDNA samples for each tissue were diluted in water to 20 ng/ μl and stored at -20°C.

Libraries description

The following five cDNA libraries were prepared from a mixture of mRNAs from winter wheat Norstar: 1) aerial parts (leaf and crown) from control and long cold acclimated wheat; 2) root tissue from control, cold-acclimated and salt stressed wheat; 3) aerial parts of wheat during dehydration stress; 4) crown from vernalized and different developmental stages of spike formation in wheat; 5) 50% crown tissues

and 50% leaf tissues from wheat after short exposures to low temperature in the light and in the dark. For cDNA synthesis, the superscript plasmid system from Gateway technology and cloning kit (Invitrogen) was used, except that the precipitation steps without yeast carrier were replaced by the Qiaquick PCR purification procedure (Qiagen). cDNAs were directionally cloned into the pCMV.SPORT6 vector with the SALI adaptor (GTCGACCCACGCGTCCG) on the 5' end and NotI primer adaptor (GCGGCCGCCC(T₁₅)) on the 3' end. For each library, six million primary transformants were amplified and frozen as glycerol stocks.

Data mining

CBF protein sequences from *Arabidopsis* were used for data mining of the NCBI and FGAS (Functional Genomics of Abiotic Stress) Databases for related wheat EST sequences. ESTs were assembled into virtual mRNAs using CAP3 ([HTTP//bio.infom-firc.it/assembly/.html](http://bio.infom-firc.it/assembly/.html)) with the two following parameters: 40 bases overlap length and 97 overlap percent identity. The assembly produced 35 contigs (virtual mRNAs containing at least 2 ESTs). Different approaches were used to isolate CBF genes from wheat. Probes corresponding either to the AP2 DNA-binding region or incomplete CBF sequences were used to screen two wheat cDNA libraries (CAP GATE 1 and library 6) using standard molecular biology techniques.

Because this approach is time consuming and costly, two others approaches were tested. First, the FGAS wheat EST database was screened for the AP2 DNA binding region. Second, public databases (NCBI, TIGR, ...) were screened with the conserved AP2 region and the ESTs present were collected and assembled into consensus sequences (contigs) using the CAP EST Assembler program (<http://bio.ifom-firc.it/ASSEMBLY/assemble.html>). A total of 35 contigs were obtained and specific primers were designed in the 5' or 3' non coding regions of contigs absent in the FGAS ESTs collection. Amplifications were performed using

PFX polymerase according to the manufacturer's instructions (Invitrogen). All positive clones from these approaches were cloned in *Escherichia coli* TOP 10 and sequenced (using the CEQTM 2000 DNA Analysis System (Beckman) at McGill University).

Northern blots and mapping by PCR

Genomic DNA was extracted by the cetyl-trimethyl-ammonium bromide method (Limin et al., 1997) from several wheat genetic stocks (cv Chinese Spring): nullisomic-tetrasomic lines (a pair of chromosomes is substituted for another pair of homeologous chromosomes) and deletion lines for homeologous group 5AL and 5DL chromosomes (Endo, 1988; Endo and Gill, 1996) generated using the chromosome of *Aegilops cylindrical*. For northern blot analyses, total RNA was extracted from wheat leaves as described (Frenette-Charron et al., 2002), and equal amounts (10µg) were electrophoresed on formaldehyde gels. Transfers and hybridizations were performed as previously described (Houde et al., 1992). Specific probes for the 15 *TaCBF* cDNAs were designed outside of the AP2 domain, to avoid possible cross-hybridisation of the probe to the transcripts of other AP2-domain containing genes. Primers used for amplification are listed in annexe 4. All filters were washed at high stringency (0.1 X SSC, 0.1 % SDS, 65°C), then they were exposed to X-ray films (BioMax-MS, Eastman Kodak, Rochester, NY) or to phosphor screens. The films were developed and the phosphor screens scanned using a Molecular Imager FX (Bio-Rad). Quantification was performed using the Quantity One software, version 4.2.2 (Bio-Rad).

Real-Time quantitative RT-PCR

Gene-specific fluorescent TaqMan-MGB and non-fluorescent primers

A detailed analysis of the nucleotide sequences of the 15 *TaCBF* genes was performed, considering that in hexaploid wheat 3 copies of each gene are expressed.

Thus gene-specific regions for each copy were identified in order to monitor the expression of each individual gene. BLASTN searches against dbEST and nr were performed to confirm the gene specificity of the nucleotide sequences chosen for the primers. The fluorescent TaqMan-MGB probes as well as the non-fluorescent primers (Table III) were designed using a combination of software, primer express 2.0 from Applied Biosystems and Primer3. Non-fluorescent primers were synthesized by Invitrogen and TaqMan-MGB probes by Applied Biosystems.

PCR amplification

Real-Time-PCR assays for each cDNA target were performed in quadruplicate on an ABI Prism 7000 sequence detection system (Applied Biosystems) using the endogenous reference 18S ribosomal protein (Applied Biosystems # 4319413E) as internal standard. The cDNA (2 μ L; 40 ng) of the RT reaction was used as a template in a 25 μ L PCR reaction containing 1x TaqMan universal PCR master mix, 0.9 μ M non-fluorescent primers, 0.25 μ M TaqMan-MGB fluorescent probe. The PCR thermal cycling parameters were 50°C for 2 min, 95°C for 10 min followed by 50 cycles of 95°C for 15 s and 60°C for 1 min.

Data analysis

All calculations and statistical analysis were performed by the SDS RQ Manager 1.1 software with a RQ min / max confidence set at 95 %. Amplification efficiency (90% to 100%) for the 13 primer sets was determined by amplification of cDNA dilution series using 80, 20, 10, 5, 2.5, and 1.25 ng per reaction (data not shown). Specificity of the RT-PCR products, as determined by gel electrophoresis, resulted in a single product with the expected length.

Phylogenetic analysis

Sequences used in this analysis are presented in annexe 1. The clustalX v-1.83 software (Thompson *et al.*, 1997) was used to generate the sequence alignment with the following parameters: gap opening penalty of 15.0, gap extension penalty of 0.30, and substitution Gonnet scoring matrices. All NJ tree was generated using TREECON for Windows (version 1.3b) with 100 bootstrap replicates and with the distance calculation set to the Poisson correction (van de Peer and de Wachter, 1994).

Results

Identification and classification of *TaCBF*

The CBF family members are known as important regulators of cold tolerance in plants. However, the identification, characterization and function of this family in cereals, compared to *Arabidopsis*, are still poorly understood. Toward this goal, we began the identification of all CBF members from wheat. Our strategy (see experimental procedures) allowed the identification of 37 cDNAs encoding putative *TaCBF* proteins (tableau 1). Analyses of these proteins using the alignment with CLUSTAL X and neighbor-joining algorithm with the wheat and rice CBF proteins distinguish 15 different *TaCBF* members (figure1). To be consistent with the established *TaCBF1* to *TaCBF3* nomenclature, we assigned consecutive number to each novel wheat CBF form, thus designating the families *TaCBF1* through *TaCBF15*. Similar but distinct genes that display high similarity at the nucleotide level along their entire length were assigned as members of a gene subfamily and given the same number and sequential numbers to reflect this close relationship. The sequences and characteristics of the 15 ORFs are listed in annexe 1. The identified *TaCBFs* genes encode a variety of proteins range from 202 to 290 amino acids. The similarity between the sequences of the 15 *TaCBFs* ranges from 65 to 85% and 63-68% identical to those of the *Arabidopsis* CBF1 to CBF6 (annexe 3).

Phylogeny Analysis

The monocot CBFs share a number of common contextual characteristics with dicot CBFs (Figure 3). Like dicot CBFs, all reported monocot genomic CBF sequences lack introns. The primary protein structure is likewise conserved and has the following general amino to carboxy terminal major features: a sequence of approximately 15–40 amino acids (I), an AP2 DNA binding domain directly flanked

by the conserved CBF subfamily signature motifs (AP2 domain), and an acidic C-terminal domain (II) postulated to act as a transcriptional activation region (Figure 3). The region I is variable in overall length and composition, and similarity is only observed among closely related CBFs. The wheat CBFs have an overall acidic character in general, which is shared with their monocot homolog.

Figure 3 shows an alignment of the highly conserved AP2 DNA binding domain and flanking CBF signature motifs for the 6 *Arabidopsis* CBF genes, and the 15 *TaCBF* genes described in this work, that contain the signature motifs. The AP2 domain of each monocot CBF (size range of 58–68 amino acids) is highly conserved, the first signature motif is also highly conserved (Figure 3). One common variation observed in this motif between *TaCBF2* and *TaCBF14* is the occurrence at the fourth amino acid position of an arginine (R) in place of the more common proline. This variation is also observed in a wide range of dicot CBFs like *ArCBF4*, *ArCBF5* and *ArCBF6* (figure 3) (Jaglo et al., 2001). The others CBFs utilizing the proline form are present in all the subgroups (figure 2). In monocots, the DSAWR motif of Jaglo et al. (2001) is more variable and the degree of divergence from the consensus correlates with some of the *TaCBF*-subgroup membership (Figure 3). Subgroup-3 members tend to contain a consensus DSAWR motif, subgroup-1 members are typically missing the C-terminal arginine residue, and subgroup-2 and 4 members generally contain the most diverged motif form (Figures 3). Three short contextual features are typically present in the acidic C-terminal domain (Figure 3): the majority of the monocot CBFs tend to either terminate with the residues LWS (Y), or have this sequence very close to their C-terminal end, and within fifteen residues upstream of this sequence, a small clustered block of 3–5 acidic amino acids is typically observed (Figure 3). However, a lot of variable are present among the residues LWS (Y) in the 15 *TaCBFs*.

Many studies have showed that monocot and dicot CBF sequences are separated on phylogenetic tree (Xiong and Fei 2006) suggesting that specialization of present day CBF gene functions have evolved independently in plants. Phylogenetic analyses have been undertaken with the aim of understanding the evolutionary relationship between CBF proteins from wheat, rice and *Arabidopsis*. The Phylogenetic tree of wheat CBFs revealed four major phylogenetic subgroups and all the rice CBFs fall in one of the four subgroups (Figure 2). The fifth subgroup contains all of the *Arabidopsis* CBFs. We designated the five subgroups 1 through 5. The 15 wheat CBFs are relatively well distributed between the four groups. Subgroup 1 through 4 are clearly defined phylogenetically with a single defined root branch each. The phylogenetic analysis showed that subgroup 4 is ancestor to the first three subgroups and was present before the divergence of these subgroups. Subgroup 4 is composed of two distinct gene family clusters, the *TaCBF11* and the *TaCBF14* families (Figure 2). The results also suggest that while an increased CBF family size is consistent between these two cereals (rice and wheat), one LT-tolerant and one LT-sensitive, the duplication and divergence events leading to increased subgroup 4 complexity may have occurred after separation of LT-tolerant cereals from the LT-sensitive member. The substantial number of *CBF* genes in monocots (15 in wheat and 10 in rice so far identified) compared to *Arabidopsis* (only six genes) suggests gene duplication and/or divergence during the evolution of angiosperms. This observation suggests that all CBF genes, at least from the angiosperms, originate from a common ancestor 60 million years later. It seems likely that this ancestor lived in a cold environment that contributed to freezing tolerance.

Mapping of *TaCBF* genes

Using nullisomic-tetrasomic lines, 13 *TaCBFs* were mapped to group 5 chromosomes and one (*TaCBF14*) to chromosome 2 (data not shown). Deletion lines for homologous group 5 chromosomes from cv Chinese Spring were used to better delimit their location. Data obtained confirmed that *TaCBFs* are located on the long arm of chromosome 5, more precisely between breakpoints 5AL-10 and 5AL-12, 5BL-1 and 5BL-16, and 5DL-1 and 5DL-2 (Figure 4). The *TaCBF14* gene is located on chromosome 2AL. unfortunately, we were not able to locate *TaCBF11* gene.

Expression profiling of *TaCBFs*

Northern blot analyses were first used to determine the expression pattern of the 15 *TaCBFs* transcripts. The results in figures 5-7 show the expressions patterns for ten genes under normal and LT conditions. For 5 *TaCBFs* (*TaCBF-3*, -6, -10, -14 and -15) no signal was detected. To facilitate the interpretation of these data, expressions patterns were classified into 3 classes according to the genotype, the level and the duration of expression under LT treatment. The first depicts a pattern where the transcript accumulate in winter wheat under LT treatment after 1h, reach a maximum level at 4-6h and then decline at very low level at 12h (*TaCBF-2*, -4, -9, -12 and -13) (Figure 5). The second shows the transcript pattern that accumulates in both genotypes, 1h following the exposure to 4°C, their expression reaches a maximum level at 4-6h, and declined to 2 fold level up to 7 days of exposure to 4°C (*TaCBF-1*, -5, -7 and -8) (Figure 6). The third, (*TaCBF-11*) represents transcript that accumulate in spring wheat 30 minutes following exposure to 4°C and stay up regulated at a maximum level up to 7 days (Figure 7). Further northern analyses show no significant effect of several abiotic stresses (cold, drought, heat shock and salt) on the expression of all *TaCBFs* (data not shown). Overall, these studies show that most

TaCBFs transcriptional levels are higher in winter wheat compared with the less freezing tolerant spring wheat.

To precisely quantify the differential RNA levels of the 15 *TaCBF* genes, we performed quantitative RT-PCR analyses that specifically amplify each *TaCBF* copy. TaqMan specific primers were designed and the sensitivity of the qRT-PCR allowed us to monitor the expression level of the 15 *TaCBF* genes. All transcripts are up-regulated by LT reaching a maximum level 2h following LT exposure (Figure 8). For convenience, the same criteria for northern blot analyses were used to categorize gene expression. One pattern confirmed that the transcript preferentially accumulate (3500 fold) in winter wheat early following LT treatment (*TaCBF*-2, -4, -6, -9, -10, -12 and -15) (Figure 8, group1). Another pattern reveals the accumulation of the transcript in both winter and spring wheat (*TaCBF*-1, -3, -7, -8, and *TaCBF*-13) (Figure 8, group2). The third pattern was observed with the *TaCBF*-5, -11 and -14 genes and show that the transcripts accumulate (6 fold) mainly in the less freezing tolerant spring wheat cultivar Manitou (Figure 8, group3).

Discussion

The CBF cold-responsive pathway is a major regulon of cold acclimation in plants. We determined that wheat has a large and complex CBF family which is representative of cereals, and demonstrated that wheat CBFs share functional characteristics with dicot CBFs, including (i) the general primary domain structure, (ii) transcript accumulation in response to abiotic stress and, in particular, cold stress, (iii) induction of *cor* gene expression. Together, these traits help to validate assignment of the wheat AP2-domain genes (and the closely related other monocot genes) described in this study as members of the plant CBF family.

In this study, the cultivar Norstar, a LT tolerant wheat genotype that has been used as a parent in a number of winter hardiness trait mapping populations, was used to isolate the CBF genes. We were able to isolate the Norstar alleles of the 37 *TaCBF* genes, demonstrating that a single genotype harbors nearly all these genes.

We found that rice, a LT-sensitive cereal, and Barley, a LT-tolerant cereal, also have expanded CBF family sizes. Likewise, the CBF family complexity was representative of cereals, as the monocot CBFs (wheat and rice) grouped into one of the four phylogenetic subgroups (Figure 2). These results suggest that an expanded CBF family and establishment of the four subgroups occurred prior to cereal speciation. While the dicot and monocot CBFs share a similar primary domain structure, the only extensively conserved regions at the sequence level are the AP2 domain and flanking CBF signature motifs. The signature motifs of Jaglo et al. (2001), conserved between monocots and dicot CBFs, were useful in distinguishing monocot CBFs from other AP2-domain family members. Common monocot variations to the CBF signature motifs were noted based on *TaCBF*-subgroup membership and the updated motifs can be defined for subgroup-1, -2, -3 and -4

members, respectively, as xKRPAGRTKFRETRHP vs AKRPAGRTKFRETRHP vs PKRPAGRT/IKxxR/KETRHP vs PKRR/PAGRTKV/FRETRHP and DSAW- vs DSAE vs DSAWR vs DSA/P. The subgroup-4 consensus pair was the most similar to the dicot motifs (Arabidopsis). When we compare with the HvCBF motif of Skinner *et al* (2005) for barley, the motifs are well conserved. The function of these conserved motifs is unknown, although the first motif displays characteristics of a nuclear localization signal (Stockinger et al., 1997). The common acidic character of the C-terminal domain is contextual, as an extended conserved region is not even present between the monocot CBFs. Recent work suggests hydrophobic amino acid side chain characteristics may be more important for trans-activation though (Wang et al., 2005).

The distribution of the 15 *CBFs* in wheat and the genes 33637700_ from rice shows that *TaCBF11* and *TaCBF14* seem to be the most ancient members from monocots. As descendents of the common ancestor invaded different environments, changes in selection pressure could have induced variants of the CBF pathway (Zhang et al, 2004). The analysis also suggests that several gene duplication events followed by selection or divergence may have occurred. Such events somehow indicate the degree to which the recruitment of new or more complex functions has occurred in this genes family. They are *a priori* at the base of the evolutionary adaptation of wheat to abiotic stresses such as freezing temperatures (Zhang et al, 2004).

Mapping studies indicated that *TaCBFs* might be organized in tandem on chromosome 5 (except *TaCBF-14* located on chromosome 2). The *Triticeae* group 5 chromosomes contain at least two closely positioned major QTLs for winter hardiness and for freezing tolerance (*Fr-A1* and *Fr-A2*) that regulate the transcription of cold regulated genes (*COR*) (Pan et al., 1994; Galiba et al, 1995). Moreover, our mapping data place *TaCBFs* in the same region as the *Rcg1*. *Rcg1* locus has an effect

on the threshold temperature at which *cor14b* and presumably other low temperature-induced genes (Choi et al., 2002). These may suggest the possibility that *TaCBFs* may be the same locus as *Rcg1* and that the difference in CBF transcript level are directly regulated at the FR-2 locus. The co-localization of the CBF genomic sequences and of the Fr A2 locus (Miller et al., 2005) imply the presence of polymorphisms in the regulatory region of some CBF gene and or the presence of sequence involved in the regulation of some cold-induced CBF in this region. If that is the case, *TaCBFs* genes can be the determinants of freezing tolerance. More extensive genomic sequences than the ones currently available (Miller et al. 2005) would be required to compare the promoters from different alleles.

Surprisingly, *TaCBF14* was mapped on the chromosome 2. Since no winter hardiness QTL has been reported in this region, this location may suggest that this gene is a pseudo gene. However, *TaCBF 14* expression was detected in the less freezing tolerant spring wheat cultivars Manitou under cold condition and HvCBF 1 in barley, its closest homolog, is induced by cold and drought (skinner *et al*, 2005). Costa et al (2001) and Qi et al (1998) have observed segregation distortion on 2H and 6H in barley mapping population, but in chromosome 5H, no segregation distortion was observed (Francia et al, 2003).

We have identified 15 *TaCBFs* genes in wheat and most are up regulated during exposition to 4°C. In rice, 10 *CBF* genes were identified but only 3 have been reported to be induced by cold (Lee et al, 2004). In Arabidopsis, 3 *CBF* genes out of 6 are cold-induced and in tomato only 1 *CBF* gene out of 3 is induced by cold (Zhang et al, 2004). Although Arabidopsis and rice both have 3 cold-induced *CBFs* genes in their genomes, the difference in their FT is undisputable (Arabidopsis LT₅₀ -9°C and rice LT₅₀ 4°C) (Uemura et al, 2003). So, the number of CBF genes present in genomes may not be as important as we would think, but the specificity of their functions might be. However, when we look at our results, where we compare the

winter cultivar and the spring cultivar, 2 *TaCBFs* (14 and 5) are not up regulated in the winter cultivar and 5 *TaCBFs* (2, 6, 9, 12 and 15) in the spring cultivar. Therefore, we can suggest that in some cases, the number as well as the function specificity of these CBFs might dictate the FT capacity.

The difference between winter and spring cultivar has been closely associated with the accumulation of COR proteins (Thomashow, 2000; Fowler et al., 1996). Winter wheat is more freezing-tolerant ($LT_{50}=-20^{\circ}\text{C}$) than spring wheat ($LT_{50}=-8^{\circ}\text{C}$) (Danyluk et al, 1996). Tolerance to low temperature is necessary for winter-habit cultivars grown in areas with subzero winter temperatures and many researches showed that FT in wheat, especially winter cultivars, is much higher than species like *Arabidopsis* (Miller et al., 2005, Vagujfalvi et al., 2005). In fact, the highest level of accumulation in winter wheat was observed for *TaCBF9* (3500fold), while in spring wheat the highest accumulation was observed for *TaCBF5* (1800fold). The higher level of expression of cold-induced genes in the winter cultivars seems to contribute to its higher cold tolerance.

The diversity in the expression patterns of these 15 *TaCBFs* genes appears to be expected and essential for cold-acclimation response. Jaglo et al (2001) showed that wheat possesses several CBF transcripts that differ in size, and that the different transcripts appear at different times. It has been shown in *Arabidopsis* that different CBFs show differences in their spatial and temporal expression (Gilmour et al., 1998; Sakuma et al., 2002). This supports the observation that each *TaCBF* must have specific target genes, and that those genes may be responsible for the difference between the tolerant and the non-tolerant plants. The fact that rice is much more sensitive to cold than *Arabidopsis* is probably due to the lack of rice homologues to some *Arabidopsis* CBF/DREB1 target genes. Dubouzet et al (2003) showed that although the CBFs in rice are functionally similar to *Arabidopsis CBF*, the specificity of their binding to CRT/DRE element is somewhat different. Such slight differences

in the recognition sequences may influence the expression profile of each CBF/DREB target gene (Sakuma et al, 2002). This again supports our hypothesis that the number as well as the function specificity of these *TaCBFs* dictates the FT capacity.

Cold tolerance is the result of complex mechanisms, including changes in membrane lipid composition, accumulation of compatible solutes such as sucrose and proline, and the expression of COR genes (Thomashow, 1999). The transcript levels of *TaCBFs* increase within 1 h of transferring plants to 4°C followed by accumulation of COR gene transcripts. These results indicated that *TaCBFs* are transcriptional activators that can activate CRT/DRE-containing genes and thus were the probable regulators of COR gene expression in wheat. The low temperature pathway in *Arabidopsis* is also characterized by the rapid induction of *CBF* genes, followed by expression of the *CBF* regulon, which includes genes (*RD29* or *COR15*) that increase plants freezing tolerance. In addition, Jaglo-Ottosen et al (1998) have shown that constitutive overexpression of CBF1 in transgenic *Arabidopsis* plants results in expression of CRT/DRE-controlled COR genes without a low temperature stimulus. Thus, *TaCBFs* could be important regulators of the cold acclimation response, controlling the level of COR gene expression, which, in turn promote freezing tolerance.

On the bases of their sequences, the Cbf genes in the temperate cereals are significantly more diverse than the *Arabidopsis* Cbf genes (Miller et al., 2005). The three *Arabidopsis* Cbf genes are more related to the *T. monococcum* Cbf6 gene, a sequence located on the chromosome 7Am, than to the 11 Cbf genes located at the Fr-Am 2 locus (Miller et al., 2005). In our study, The six *Arabidopsis* Cbf genes are more related to the *T. aestivum* Cbf14 gene, a sequence located on the chromosome 2, than to the 13 Cbf genes located at the Fr-Am 2 locus Therefore, specific studies of the functions and redundancy of the Cbf genes present in the temperate cereals will

be necessary to clarify the roles of the wider variety of Cbf genes. The comparative analyses of Cbf expression in a number of wheat, barley and triticum monococcum genotypes characterized by different level of frost resistance might provide a new insight on the role of Cbf genes in the regulation of frost resistance. If CBF proteins have redundant functional activities, frost resistance might be found associated to the overexpression of different members of the Cbf gene family. On the contrary, if frost resistance is strictly associated to the over-expression of some CBF, these genes might be more important than other Cbfs in the achievement of frost resistance.

In conclusion, our data support the importance of the *TaCBFs* genes on group 5 chromosome that may fulfill similar functions in controlling the level of low temperature-regulated gene expression and promoting freezing tolerance. The discovery of the CBF cold-response pathway and demonstration that the expression of CBF genes increases freezing and chilling tolerance provide the tools for improving the stress tolerance of agriculturally important plant through optimizing the expression of the CBF genes. Considering the large number of genes induced in response to low temperature, we hypothesized that differences in the sequences of the CCGAC core element and/or in the sequences that surround it may result in the recruitment of distinct CBF proteins. Which, support the observation that the differences in the expression patterns between CBF genes in Triticeae genomes may regulate different low-temperature-regulated genes, which should have separate and distinct signal transduction pathways toward the activation of freezing tolerance. However, the evolutionary scenarios presented above are highly speculative and represent only a few of many possibilities. Indeed, developing scenarios regarding the evolution of the CBF cold response pathway is complicated by the fact that the CBF regulon of genes can have multiple functions. In *Arabidopsis*, for instance, the CBF regulon not only imparts freezing tolerance, but also imparts tolerance to chilling temperatures, dehydration stress and high salinity. Thus, selective pressures for

functional CRT/DRE elements and the number and regulon of CBF genes would not be expected to be limited to freezing temperature, but involve other stresses as well.

References

- Bannai, H., Tamada, Y., Maruyama, O., Nakai, K., and Miyano, S. (2002) Extensive feature detection of N-terminal protein sorting signals. *Bioinformatic*. 18(2): 298-305
- Choi, D.W., Rodriguez, E.M. and Close, T.J. (2002) Barley *CBF3* gene identification, expression pattern, and map location. *Plant Physiol*. 129, 1781-1787.
- Costa, M.A., Silva, N.C., and Castro-Prado, M.A. (2001) Genetic and cytological characterization of a developmental mutant of *Aspergillus nidulans* induced by 5-azacytidine. *Biol Res*. 34, 91-98.
- Couderchet, M. and Koukkari, W.L. (1987) Cold sensitivity oscillations of young soybean plants. *Prog Clin Biol Res* 227A: 59-65.
- Danyluk, J., Carpentier, E. and Sarhan F. (1996) Identification and characterisation of a low temperature regulated gene encoding an actin-binding protein from wheat. *FEBS Lett*. 389, 324-327.
- Dubcovsky, J., Lijavetzky, D., Appendino, L., and Tranquilli, G. (1998). Comparative RFLP Mapping of *Triticum monococcum* genes controlling vernalization requirement. *Theor. Appl. Genet*. 97, 968-975.
- Dubouzet, J.G., Sakuma, Y., Ito, Y., Kasuga, M., Dubouzet, E.G., Miura, S., Seki, M., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2003) OsDREB genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. *Plant J*. 33: 751-763.
- Endo, T.R. (1988). Induction of chromosomal structural changes by a chromosome of *Aegilops cylindrica* L. in common wheat. *J. Hered*. 79, 366-370.
- Endo, T.R., and Gill, B.S. (1996). The deletions stocks of common wheat. *J. Heredity* 87, 295-307.
- Feng, G., Zhang, F., Li, X., Tian, C., Tang, C. and Rengel Z. (2002) Improved tolerance of maize plants to salt stress by arbuscular mycorrhiza is related to higher accumulation of soluble sugars in roots Mycorrhiza. 12:185-90.

Fowler, S.G., Cook, D. and Thomashow, M.F., (2005) Low temperature induction of *Arabidopsis* *CBF1*, 2, and 3 is gated by the circadian clock. *Plant Physiol.* 137: 361–368.

Fowler, B., Sarhan, F., Limin, A., and Chauvin, L.P. (1996) The regulatory role of vernalization in the expression of low-temperature induced genes in wheat and rye. *Theor.Appl.Genet.* 93: 554-559.

Francia, E., Rizza, F., Cattivelli, L., Stanca, A. M., Galiba, G., Tóth, B., Hayes, P. M., Skinner, J. S. and Pecchioni, N. (2003) Two loci on chromosome 5H determine low-temperature tolerance in a 'Nure' (winter) x 'Tremois' (spring) barley map. *Theor Appl Genet.* 108: 670-80.

Frenette Charron, J.B., Breton, G., Badawi, M., and Sarhan, F. (2002) Molecular and structural analyses of a novel temperature stress-induced lipocalin from wheat and *Arabidopsis*. *FEBS Lett.* 517(1-3): 129-32

Galiba, G., Quarrie, S.A., Sutka, J., and Morgounov, A. (1995). RFLP Mapping of the vernalization (*Vrn1*) and frost resistance (*Fr1*) genes on chromosome 5A of wheat. *Theor. Appl. Genet.* 90: 1174-1179.

Gilmour, S. J., Zarke, D. G., Stockinger, E. J., Salazar, M. P., Houghton, J. M., and Thomashow, M.F. (1998) Low temperature regulation of the *Arabidopsis* CBF family of AP2 transcriptional activators as an early step in cold-induced COR gene expression . *Plant J.* 16: 433-442.

Gilmour, S.J., Sebolt, A.M., Salazar, M.P., Everard, J.D. and Thomashow, M.F. (2000) Overexpression of the *Arabidopsis* CBF3 transcriptional activator mimics multiple biochemical changes associated with cold acclimation. *Plant Physiol.* 124: 1854-1865.

Gilmour, S.J., Fowler, S.G., and Thomashow, M.F. (2004) *Arabidopsis* transcriptional activators CBF1, CBF2, and CBF3 have matching functional activities. *Plant Mol Biol.* 54: 767-81.

Gong, Z., Lee, H., Xiong, L., Jagendorf, A., Stevenson, B. and Zhu, J.K. (2002) RNA helicase-like protein as an early regulator of transcription factors for plant chilling and freezing tolerance. *Proc. Natl Acad. Sci. USA*, 99:11507-11512.

Guy, C.L. (1990) Cold acclimation and freezing stress tolerance: role of protein metabolism. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 41:187-223.

Haake, V., Cook, D., Riechmann, J.L., Pineda, O., Thomashow, M.F. and Zhang, J.Z. (2002) Transcription factor CBF4 is a regulator of drought adaptation in *Arabidopsis*. *Plant Physiol.* 130: 639-648.

Jaglo, K.R., Kleff, S., Amundsen, K.L., Zhang, X., Haake, V., Zhang, J.Z., Deits, T., and Thomashow, M.F. (2001) Components of the *Arabidopsis* C-repeat/dehydration-responsive element binding factor cold-response pathway are conserved in *Brassica napus* and other plant species. *Plant Physiol.* 127:910-917.

Jaglo-Ottosen, K.R., Gilmour, S.J., Zarka, D.G., Schabenberger, O., and Thomashow, M.F. (1998) *Arabidopsis* CBF1 overexpression induces COR genes and enhances freezing tolerance. *Science*. 280 (5360): 104-6

Lee, S., Huh, K., An, K., An, G., and Kim, S., (2004) Ectopic expression of a Cold-inducible Transcription Factor, CBF1/DREB1b, in Transgenic Rice (*Oryza sativa* L.). *Mol. Cells*. 18: 107-114.

Jofuku, K. D., den Boer, B. G., Van Montagu, M., and Okamuro, J. K. (1994) Control of *Arabidopsis* flower and seed development by the homeotic gene APETALA2. *Plant Cell*. 6: 1211-25.

Kasuga, M., Liu, Q., Miura, S., Yamaguchi-Shinozaki, K. and Shinozaki, K. (1999) Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. *Nat. Biotechnol.* 17: 287-291.

Limin, A.E., and Fowler, D.B. (2002). Developmental traits affecting low-temperature tolerance response in near-isogenic lines for the vernalization locus Vrn-A1 in wheat (*Triticum aestivum* L. em Thell). *Ann. Bot.* 89, 579-585.

Liu, Q., Kasuga, M., Sakuma, Y., Abe, H., Miura, S., Yamaguchi-Shinozaki, K. and Shinozaki, K. (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in *Arabidopsis*. *Plant Cell*, 10:1391-1406.

Magnani, E., Sjolander, K., and Hake, S. (2004) From Endonucleases to Transcription Factors: Evolution of the AP2 DNA Binding Domain in plants. *Plant Cell Prev.*

Medina, J., Bargues, M., Terol, J., Perez-Alonso, M. and Salinas, J., (1999) The Arabidopsis CBF gene family is composed of three genes encoding AP2 domain-containing proteins whose expression is regulated by low temperature but not by abscisic acid or dehydration. *Plant Physiol.* 119: 463-469.

Miller, A.K., Galiba G., and Dubcovsky J. (2005) A cluster of 11 CBF transcription factors is located at the frost tolerance locus FR-A^{m2} in Triticum monococcum. *Mol Gen Genomics* 275: 193-203

Nakai, K., and Kanehisa, M. (1992) A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics.* 14(4): 897-911.

Nielsen, H., and Krogh, A. (1998) Prediction of signal peptides and signal anchors by a hidden Markov model. *Proc Int Conf Intell Syst Mol Biol.* 6:122-30

Ouellet, F., Vazquez-Tello, A., and Sarhan, F. (1998) The wheat wcs120 promoter is cold-inducible in both monocotyledonous and dicotyledonous species. *FEBS Lett.* 423: 324-8.

Okamuro, J.K., Caster, B., Villarroel, R., Van Montagu, M., and Jofuku, K. D. (1997) The AP2 domain of APETALA2 defines a large new family of DNA binding proteins in Arabidopsis. *Proc Natl Acad Sci U S A.* 94: 7076-81.

Riechmann, J.L., and Meyerowitz, E.M. (1998) The AP2/EREBP family of plant transcription factors. *Biol. Chem.* 379: 633-646.

Rikin, A., Dillwith, J.W., Bergman, D.K. (1993) Correlation between the circadian rhythm of resistance to extreme temperatures and changes in fatty acid composition in cotton seedlings. *Plant Physiol* 101:31-36.

Sakai, A. and Larcher, W. (1987) Frost Survival of Plants: Responses and Adaptation to Freezing Stress. Berlin: Springer-Verlag

Sakuma, Y., Liu, Q., Dubouzet, J.G., Abe, H., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2002) DNA-binding specificity of the ERE/AP2 domain of *Arabidopsis* DREBs, transcription factors involved in dehydration and cold-induced gene expression. *Biochem. Biophys. Res. Commun.* 290: 998-1009.

Seki, M., Narusaka, M., Abe, H., Kasuga, M., Yamaguchi-Shinozaki, K., Carninci, P., Hayashizaki, Y., and Shinizaki, K. (2001) Monitoring the expression pattern of 1300 *Arabidopsis* genes under drought and cold stresses by using a full-length cDNA microarray. *Plant Cell* 13: 61–72.

Shinozaki K and Yamaguchi-Shinozaki K (2000) Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. *Curr Opin Plant Biol* 3:217-23.

Skinner, J.S., Zitzewitz J., Szucs, P., Marquez-Cedillo, L., Filichkin, T., Amundsen, K., Stockinger, E. J., Thomashow M.F., Chen T. H.H., and Hayes P. M. (2005) Structural, functional, and phylogenetic characterization of a large CBF gene family in barley. *Plant Molecular Biology*, 59:533–551.

Stockinger, E.J., Gilmour, S.J. and Thomashow, M.F. (1997) *Arabidopsis thaliana* *CBF1* encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. *Proc. Natl Acad. Sci. USA*, 94:1035–1040.

Thomashow, M.F. (1999) Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. *Ann Rv Plant Physiol Plant. Mol. Biol.* 50: 571-599.

Thomashow, M.F. (2001) So what's new in the field of plant cold acclimation? Lots! *Plant Physiol.* 125: 89-93.

Vagujfalvi, A., Aprile, A., Miller, A., Dubcovsky, J. Delugu, G., Galiba, G., and Cattivelli, L. (2005) The expression of several CBF genes at the Fr-A2 locus is linked to frost resistance in wheat. *Mol Gen Genomics* 274: 506-514.

Vagujfalvi, A., Galiba, G., Cattivelli, L. and Dubcovsky, J. (2003) The cold-regulated transcriptional activator Cbf3 is linked to the frost-tolerance locus Fr-A2 on wheat chromosome 5A. *Mol Genet Genomics.* 269:60-67.

Vazquez-Tello, A., Ouellet, F., and Sarhan, F. (1998) Low temperature-stimulated phosphorylation regulates the binding of nuclear factors to the promoter of Wcs120, a cold-specific gene in wheat. *Mol Gen Genet.* 257: 157-66.

Yamaguchi-Shinozaki, K. and Shinozaki, K. (1994) A novel cis-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature, or high-salt stress. *Plant Cell*, 6: 251-264.

Zhang, X., Fowler, S.G., Cheng, H., Lou, Y., Rhee, S.Y., Stockinger, E.J., and Thomashow, M.F. (2004) Freezing-sensitive tomato has a functional CBF cold response pathway, but a CBF regulon that differs from that of freezing-tolerant *Arabidopsis*. *Plant J.* 39: 905-924.

Zhang, J. (2003) Evolution by gene duplication: an update. *Trends Ecol. Evol.* 18: 292-298.

Legends to figures

Tableau 1: wheat genes containing the AP2 domain.

Genes are listed under their clone ID and to be consistent with the established *TaCBF1* to *TaCBF3* nomenclature, we assigned consecutive number to each novel wheat CBF form, thus designating the families *TaCBF1* through *TaCBF15*. Similar but distinct genes that display high similarity at the nucleotide level along their entire length were assigned as members of a gene subfamily and given the same number and sequential numbers to reflect this close relationship. The table also shows their cDNA size, the number of amino acids and their location on the chromosome as well as their expression profile.

Figure 1: Phylogenetic analyses of wheat CBF proteins.

Phylogenetic analyses of 37 CBF proteins from hexaploid wheat reveal 15 different groups of maximum likelihood. The tree was constructed as described in methods using the rice protein (50253272) as an outgroup which shows the highest homology with dicot CBFs. Red squares indicate proteins that were analysed by northern blot.

Figure 2: Evolution of the CBF subgroup in plants.

Phylogenetic analyses of *TaCBF1-15* proteins from the monocotyledonous plants wheat and rice and the dicotyledonous *Arabidopsis* show the maximum likelihood. The tree showing the evolution of one representative of each of the 15 different wheat groups, the 10 genes found in the rice genome and the 6 genes found in *Arabidopsis* was constructed as described in material and methods using the *Arabidopsis* protein CBF4as an outgroup. Vertical bars denote the respective *TaCBF*-subgroups and members.

Figure 3: Structure and sequence alignment of the 15 *TaCBF*s with 6 *ArCBF*s members.

Identical residues are in black and similar residues are in gray. The ERF/AP2 DNA-binding domains are underlined by a red line. The NLS sequence (I) and DSA box (II) are indicated by blue lines. The conserved sequences in the C-terminal of these proteins are highlighted on a black background.

Figure 4: Mapping of *TaCBF* genes on the wheat chromosomes.

Schematic representation of the deletion lines used to map the *TaCBF* genes on chromosomes 2 and 5. The numbers indicate the deletion breakpoints of each line. Black boxes represent C-band markers. Red boxes indicate where the location of *TaCBF* genes.

Figure 5: Accumulation of *TaCBF2*, 4, 9, 12, 13 transcripts during cold acclimation in winter wheat cv Norstar and spring wheat cv Quantum.

A. Total RNA was extracted from wheat leaves, separated on formaldehyde agarose gel and transferred to nylon membranes. Specific *TaCBF2*, 4, 9, 12, 13 probes designed outside of the AP2 DNA binding domain were PCR-amplified. Hybridization with gene-specific 32p-labeled probes was performed. NA, non-acclimated plants grown for 7 days at 20°C; time points indicate plants exposed to cold for 15 and 30 minutes, 1, 2, 4, 6, 12 and 24 hours, and 7 days. The ethidium bromide stained 28S ribosomal band is included to show RNA loads.

B. Quantification of *TaCBF4* from group 1. Mean expression levels are presented as a proportion of the highest value after normalization with respect to 18S rRNA expression levels.

Figure 6: Accumulation of *TaCBF1*, 5, 7, 8 transcript during cold acclimation in winter wheat cv Norstar and spring wheat cv Quantum.

A. Total RNA was extracted from wheat leaves, separated on formaldehyde agarose gel and transferred to nylon membranes. Specific *TaCBF1*, 5, 7, 8 probes designed outside of the AP2 DNA binding domain were PCR-amplified. NA, non-acclimated plants grown for 7 days at 20°C; time points indicate plants exposed to cold for 15 and 30 minutes, 1, 2, 4, 6, 12 and 24 hours, and 7 days. The ethidium bromide stained 28S ribosomal band is included to show RNA loads. The ethidium bromide stained 28S ribosomal band is included to show RNA loads.

B. Quantification of *TaCBF1* from group 2. Mean expression levels are presented as a proportion of the highest value after normalization with respect to 18S rRNA expression levels.

Figure 7: Accumulation of *TaCBF11* transcript during cold acclimation in winter wheat cv Norstar and spring wheat cv Quantum.

A. Total RNA was extracted from wheat leaves, separated on formaldehyde agarose gel and transferred to nylon membranes. Specific *TaCBF11* probes designed outside of the AP2 DNA binding domain were PCR-amplified. NA, non-acclimated plants grown for 7 days at 20°C; time points indicate plants exposed to cold for 15 and 30 minutes, 1, 2, 4, 6, 12 and 24 hours, and 7 days. The ethidium bromide stained 28S ribosomal band is included to show RNA loads. The ethidium bromide stained 28S ribosomal band is included to show RNA loads.

B. Quantification of *TaCBF11* from each of the group 1. Mean expression levels are presented as a proportion of the highest value after normalization with respect to 18S rRNA expression levels.

Figure 8: Accumulation of *TaCBF3*, *9*, *11* transcript during cold acclimation in winter wheat cv Norstar and spring wheat cv Manitou starting at 20:00h.

Total RNA was isolated from leaves, reverse transcribed, and subjected to quantitative real-time PCR. Relative transcript abundance was calculated and normalized with respect to the 18S rRNA transcript level. Data shown represent mean values obtained from four independent amplification reactions. The experiment was repeated three times. Graphics of the data showing mean expression of one CBF from each of the 3 groups (CBF9=group1, CBF3=group2, CBF11=group3). The color bleu represent the genotype Norstar and the pink the genotype Manitou.

Figures

CLONE ID	cDNA size	amino acids	Expression	Mapping
TaCBF1- 1 clone9F1L7-5	1077bp	217		5A
2 L3C116-P09-23	887bp	212		5B
3 clone 9	980bp	212	cold specific	5D
TaCBF2- 1 D992	940bp	225	cold specific	5A
2 D992 second	885pb?			
3 GI25164819R1L1_7	660pb?			
TaCBF3- 1 WEF049_G17	958bp	238		
2 C25R1L1-48	975bp	242		
3 D917	959bp	238	low cold specific	5D
TaCBF4- 1 L6B005-G12	892bp	214		
2 D990	860bp	214	low cold specific	5B
3 L6B016-J14	1003bp	212		
TaCBF5- 1 clone 3-2-1	1080pb	269		5A
2 contig28-JD_2-2-7	999bp	269		5B
3 D921	1051bp	269	cold specific	5D
TaCBF6- 1 L5B016-A21-37	963bp	245	no detection	5A
2 C16R1L7-a	954pb	246	no detection	
3 L3C106-F16-	100bp	243		5B
TaCBF7- 1 L6B020-O14F1-L1-68	1059bp	290		
3 L6B020-O14F1-L1-72	1211pb	275	cold specific	5B
2 CA677056F1L1-143	1219bp	275		
TaCBF8- 1 L6B006_E20	872pb	222		
2 L6B012-G16-13	865bp	222	cold specific	5B
TaCBF9- 1 L6B002-G24-33	919bp	241	cold specific	5A
2 L6B002-G24	930bp?			

TaCBF10-	1	C133R1L1-2F1-L7-7	979pb	202	no detection	
	2	C133R1L1-4	1100bp	202		5D
TaCBF11-	1	C22R1L3-a	1000bp	228		
	2	C21R1L8-a	1024pb	219	low cold specific	?
	3	L4B201C08	907pb	225		
TaCBF12-	1	L6B007_F24-19	1000bp	234	cold specific	5D
	2	C15R1L1-20	984pb	234	low cold specific	5A
	3	L6B026-H10	954pb	234		5B
TaCBF13-	1	L6B022_I06-10	1000bp	245	cold specific	5B
	2	L6B013-003R1-L1-4	1033pb	245		
TaCBF14-	1	BQ744365R1L8-153	900bp	218	no detection	2A
TaCBF15-	1	L3C106-F16-	100bp	243		5B

Tableau 1

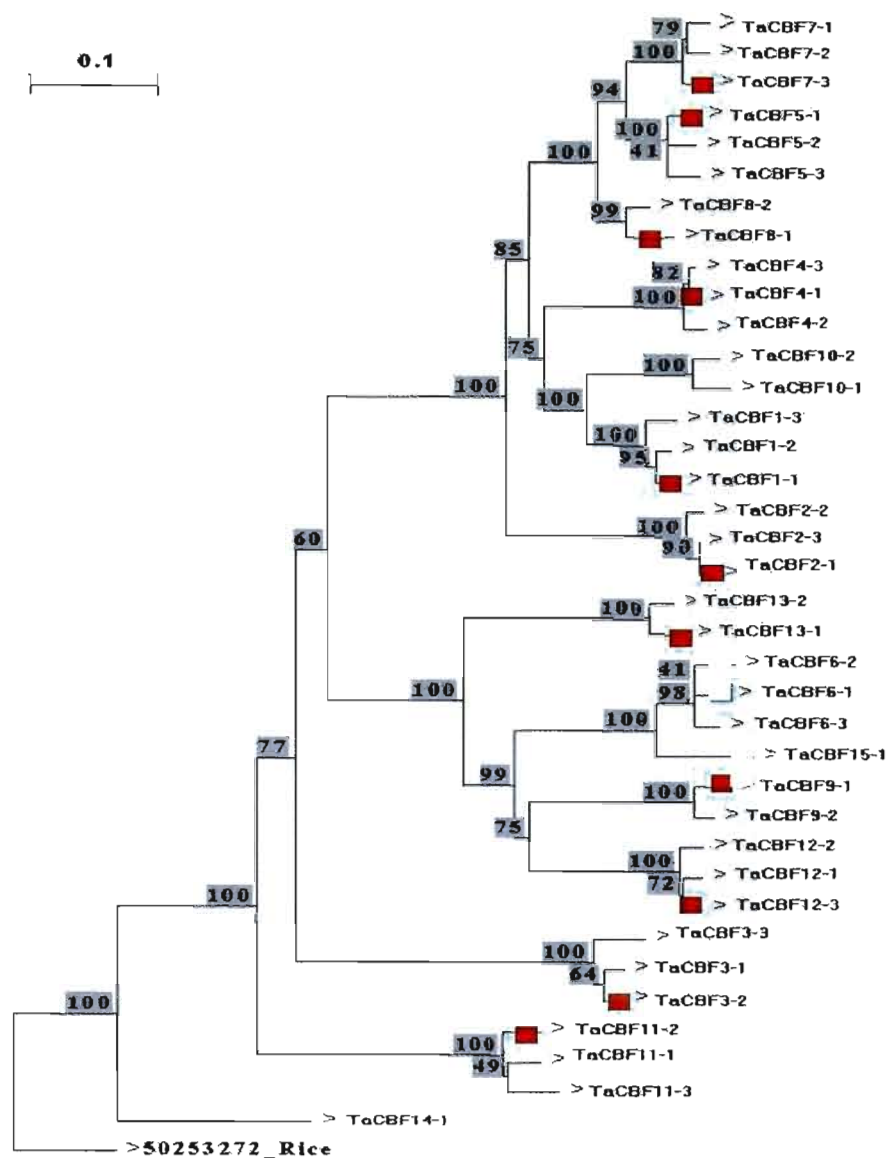


Figure 1

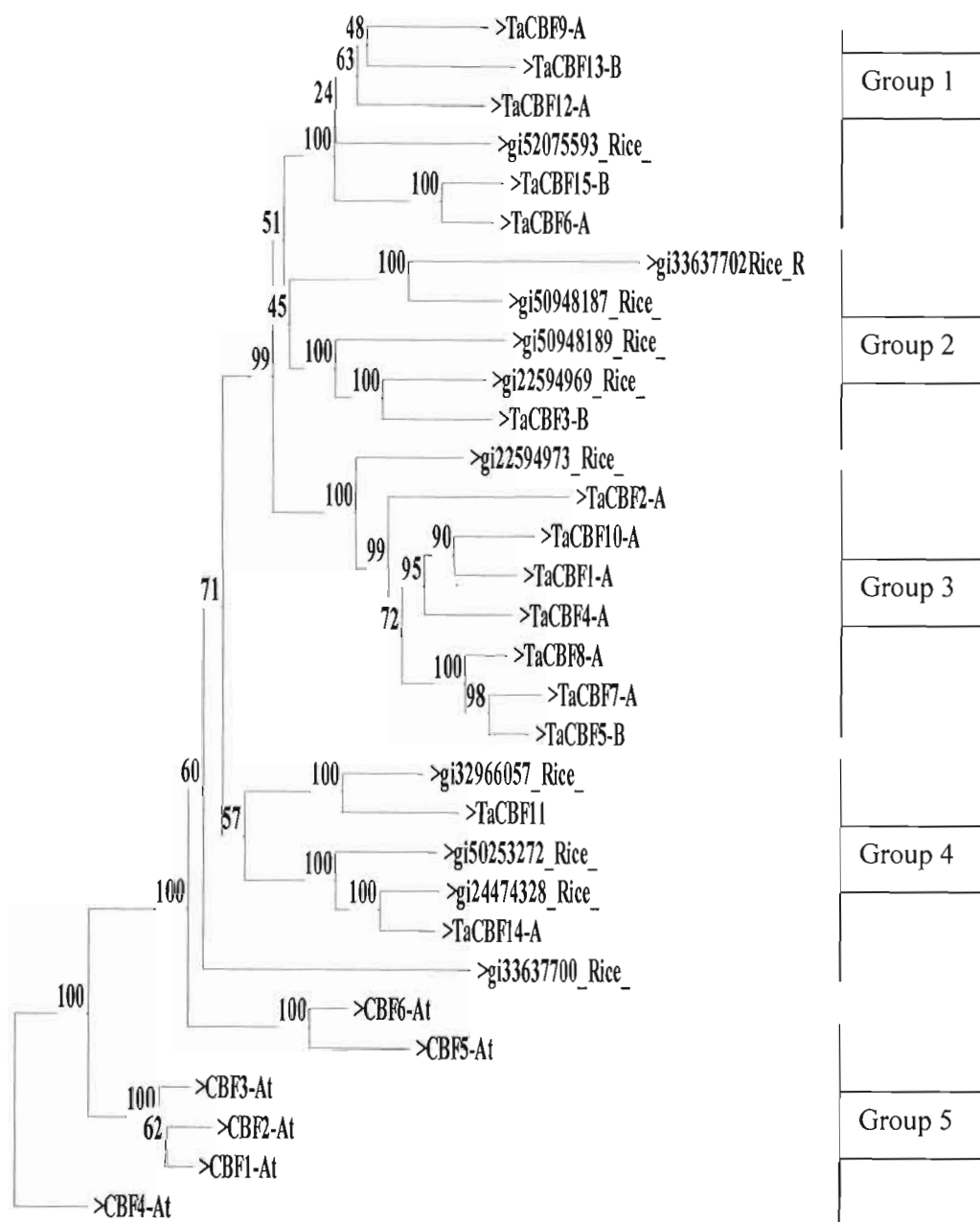


Figure 2

		I		AP2 Domain	
CBF1	1	---NN-SFSAFSEMFG--SDYEP---QGG-----DYC--PTLANSCH	EPKAGKKFRETRHI	TYRGVRR	MS-KA
CBF2	1	---NN-SFSAFSEMFG--SDYESTVSSGG-----DYS--PKLANSCH	EPKAGKKFRETRHI	TYRGVRR	MS-KA
CBF3	1	---NN-SFSAFSEMFG--SDYESTVSSGG-----DYI--PTLANSCH	EPKAGKKFRETRHI	TYRGVRR	MS-KA
CBF4	1	---NPFYSTFPDSFLSISDHRFTVSDSS-----ECS--PKLANSCH	EPKAGKKFRETRHI	TYRGVRR	MS-KA
CBF5	1	-----MN-----NDD--IILAEMRPF	KRAGRVK	KETRHVYPC	TRRI-NGDK
CBF6	1	-----ME-----NDD--ITVAEMKPF	KRAGRVK	KETRHFTYPC	VPRF-DGDK
TaCBF14	1	---EWAYSGGGHSS---SGTKSPAAGGR-----EEGSYMTVS	APKPRAGRTK	VRETRHI	VYKVESE-MP-RA
TaCBF11	1	---EYOYSYGGGDGD-----QAG-----QGG-YATVT	APKPRAGRTK	FRETRHI	LYRGVRRRQA-RA
TaCBF1	1	---MTAAAG-----SIR-----EGHRTVC	EPKPRAGRTK	FRETRHI	LYRGVRRRRL-QA
TaCBF10	1	---MAADA-----SIS-----DGHRTVW	EPKPRAGRTK	FRETRHI	LYRGVRRRGY-RA
TaCBF4	1	---MAADAA-----SIC-----DGHRTVW	EPKPRAGRTK	FRETRHI	LYRGVRRRGP-RA
TaCBF5	1	---MVADIA-----RSGQQ-----EQGHRTVS	EPKPRAGRTK	FRETRHI	LYRGVRRRGV-QA
TaCBF7	1	---MVADAA-----SISGQD-----QGHRTVS	EPKPRAGRTK	FRETRHI	LYRGVRRRGV-QA
TaCBF8	1	---MVADAA-----KSGHQ-----EQGHRTVS	EPKPRAGRTK	FRETRHI	LYRGVRRRGV-QA
TaCBF2	1	---MTNAAW-----PQFDGQEYRTVWPEEQEYRTVW	EPKPRAGRTK	FRETRHI	LYRGVRRRGV-QA
TaCBF6	1	---MT-MGLEVS-----SSSSSSSVSS-----SPVHAAGR	SLAPRAGRTK	FRETRHI	LYRGVRRRGV-QA
TaCBF15	1	---MT-MGLEVS-----SSSSSSSVSS-----SLAPRAGRTK	FRETRHI	LYRGVRRRGV-QA	
TaCBF13	1	---MT-TGPERN-----WNHASPSSL-----EQGMPTSPA	PTKPRAGRTK	FRETRHI	VYKVESE-MP-RA
TaCBF9	1	---MTMTGSDQQ-----RSQSSP-SS-----SHLAPRAGRTK	FRETRHI	LYRGVRRRGV-QA	
TaCBF12	1	---MT-FGINGW-----ISSSSSSTSGH-----ELGDAVPVW	PAAAPRAGRTK	FRETRHI	LYRGVRRRGV-QA
TaCBF3	1	MCPIKRMSGESGSPCSG--ESFYDSTSPEN-----QARQAAT	APAPRAGRTK	FRETRHI	LYRGVRRRGV-QA

		AP2 Domain		II	
CBF1	60	VSEVHE-N---KKSRWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	EST	CAKLOKAA
CBF2	63	VCHLE-N---KKSRWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	EST	CAKLOKAA
CBF3	63	VCEVHE-N---KKSRWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	EST	CAKLOKAA
CBF4	66	VCEVHE-N---KKSRWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	ETI	CPKPKKAA
CBF5	42	VCEVHE-T---HQRWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	ESN	DDPVIREV
CBF6	42	VCEVHE-I---HQRWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	EST	DDPVIREV
TaCBF14	65	VCEVHE-H---GKRWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	PGS	GHEPREEV
TaCBF11	57	VCEVHE-N---KKSRWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	EST	CAKLOKAA
TaCBF1	51	VCEVVR-R---ACVRLWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	EST	CAKLOKAA
TaCBF10	50	VCEVVR-R---TBEWLWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	EST	CAKLOKAA
TaCBF4	51	VCEVVR-L---MRSRLWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	EST	CAKLOKAA
TaCBF5	55	VCEVVR-P---IKSRLWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	EST	CAKLOKAA
TaCBF7	55	VCEVVR-P---VGSRLWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	EST	CAKLOKAA
TaCBF8	55	VCEVVR-P---VGSRLWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	EST	CAKLOKAA
TaCBF2	63	VCHLPVAGSRYSRWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	EST	CAKLOKAA
TaCBF6	61	VCEVVR-P---KSRWLWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	EST	CAKLOKAA
TaCBF15	48	VCEVVR-P---KSRWLWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	EST	CAKLOKAA
TaCBF13	62	VCEVVR-P---KSRWLWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	EST	CAKLOKAA
TaCBF9	52	VCEVVR-P---KSRWLWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	EST	CAKLOKAA
TaCBF12	62	VCEVVR-P---KSRWLWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	EST	CAKLOKAA
TaCBF3	73	VCHLPVAGSRYSRWLGTFTAEMAARAHVVAALAIRGHS	ACLNFAASA-RRI	EST	CAKLOKAA

CBF1	127	EAALAFD---ET-----C-----DTTNNHGLDMEETMV
CBF2	130	EAALAFD---EM-----C-----HMTTDAHGLDMEETLV
CBF3	130	EAALAFD---EM-----C-----DATTD-HGFDMEETLV
CBF4	133	EAALAFD---ET-----T-----TEGSK-TAAEAEAAAG
CBF5	109	EAALAFD---DL-----E-----SGITVLPACGDDVDLG
CBF6	109	EAALAFD---EF-----S-----TGITVLPASA-----
TaCBF14	133	EAALAFD---N-----N-----AATTEAPAS-PADAG
TaCBF11	126	ALAETACPAAPVA-----A-----EASAPAPAMMMQEP
TaCBF1	125	IAVLEP-RQRPVST-----EM-----HDGEKDAQGSPTPS---
TaCBF10	124	VAVVAF-RQSVAST-----A-----ADGEKDVQGSPTPS---
TaCBF4	122	VAVVAF-RQHRVASMA-----PLSPART-----TDDEKEIDGSPAPS---
TaCBF5	129	VAVVAF-RKQIIIPVAVAVVALQQQ--VPVAVAVVALQQQVPLAVAVVALQQQVQVAVAVAVVALQQQVQIIIPVAVVACLAP-
TaCBF7	129	VAVVAF-RKQIIIPVAVAVVALQQQ--VPVAVAVVALQQQVPLAVAVVALQQQVQVAVAVAVVALQQQVQIIIPVAVVACLAP-
TaCBF8	129	VAVVAF-LQQNIP-----PAACPTV-----EPTDPSNS---
TaCBF2	139	VALKER-EQQRPA-----DES-TPASSTAEES---
TaCBF6	131	EAVADP-RREAANG-SLTATVT---EEASCGAPEESS-----SESASGSETSEPSA
TaCBF15	118	EAVADP-RREAANG-SITATATATAAEASCGAPEESS-----SESASGSETSEPSA
TaCBF13	133	GAVVAF-RQREAG--ASTGAVTEVASIDG--IASAASAP-----DNASAAAASHSOPPCA
TaCBF9	122	SAVADF-RREAASGAATTSLAATVPVDDGSCSQAQSSM-----ENTGSSWTSSSSSLPSG
TaCBF12	135	QAVAG-LPEAAT-----VVPDVD-----EATSPVYLPSPVDNA
TaCBF3	145	EAVEDLLREAI-----ED-----DALSGTSSAPSLTD

```

CBF1 154 E I I T P E ---- Q S E G A F Y D E E T M F G P L L D N A E G M L P P S V Q W N H N Y ---- G E D G ---- D - V S -- L M S Y
CBF2 157 E I I T P E ---- Q S Q D A F Y D E E M I G S S L L D N A E T M L P S S V Q W N Y N F ---- V E D D ---- D - V S -- L M S Y
CBF3 156 E I I T A E ---- Q S E N A F Y H D E M F E P S L L E N A E T M L P L S V Q W N H N H ---- V D D D ---- D - V S -- L M S Y
CBF4 159 E V R E G E R A E E Q N G G V F Y D D E A L G P N F E N A E T M L P P E V G W N H N ---- F D V G ---- D - V S -- L M S Y
CBF5 137 F S G S G S G S G S - E E R N S S S Y G F G D Y P E V S T M M R L A E P L E S P I R S Y M E D M T P T N V Y T E E M C Y ---- E D M S -- L M S Y
CBF6 130 --- S - ---- E F D T S D ---- E G V A G M M M R L A E P L E S P I R S Y I D M N -- T S V Y V E E M C Y ---- E D L S -- L M S Y
TaCBF14 162 N E L V A N ---- S P Y H I M D G L E F P Q G V L D - N H G M L E P P M A G P S T ---- W I E D Y D C ---- E - V S -- L M S Y
TaCBF11 156 S V E Y D D ---- Y P M Q Y G G I G D F Q H S Y Y D G D A S G G D W Q S S H M D G A D D S N C G G Y T A G ---- E - V P -- L M S Y
TaCBF1 157 -- E L S T S --- S --- D L L D E H W F P G T A G S Y Y A S L A Q G M L E P P S A R T W S E D ---- G G E Y S A ---- V Y T P -- L M S Y
TaCBF10 153 -- E L S T S --- S --- D L L D E H W F P G T A G S Y Y A S L A Q G M L E P P S A R T W S E D ---- G G E Y S A ---- V Y T P -- L M S Y
TaCBF4 159 -- A L F M S --- S --- E L L N E H W F P G T A G S Y Y A S L A Q G M L E P P S A R T W S E D ---- G G E Y S A ---- V Y T P -- L M S Y
TaCBF5 207 -- E T Y M S --- S G D L L E L D E E Q W F P G M A G S Y Y A S L A Q G M L E P P S A R T W S E D ---- G G E Y S A ---- V Y T P -- L M S Y
TaCBF7 209 S H L Y M S --- S S D L L E L D E E Q W F P G M A G S Y Y A S L A Q G M L E P P S A R T W S E D ---- G G E Y S A ---- V Y T P -- L M S Y
TaCBF8 158 -- L Y M S --- S G D L L E L D E E H W F P G M A G S Y Y D S L A Q G M L E P P S A R T W S E D ---- G G E Y S A ---- V Y T P -- L M S Y
TaCBF2 165 -- A L S I I --- P S D L S G L D N E H W F P G M A G S Y Y A S L A Q G M L E P P S A R T W S E D ---- G G E Y S A ---- V Y T P -- L M S Y
TaCBF6 180 D E E V P --- V A V D T D M F S R D L F P E P P S Y Y A S L A E A L L D P P S T V A I I D S ---- Y W D N ---- D D G A V A -- L M S Y
TaCBF15 171 D E D A V P --- G G M D I E M F S R D L F P E P P S Y Y A S L A E A L L D P P S T V A I I D S ---- Y W D N ---- D D G A V A -- L M S Y
TaCBF13 185 N E E V P --- D A L C H D M F - E H T S E P A G Y Y A D L A Q G M L E P P P P S - S G A ---- S S E Q ---- G D A A -- L M S Y
TaCBF9 178 D E M A V P --- A T L G C N M F - E D M S E E M D Y Y A Y P E S H L E P P Q P P V - A G A ---- C W D T E ---- G S G A A A -- L M S Y
TaCBF12 169 D E V I Q V P T F S P L G S D M F - E D M S E E M D Y Y A G E A Q G M L E P P T P A - Y W E ---- T G E C ---- D G G A A A -- L M S Y
TaCBF3 176 D S S S S P --- L P E E D S P F - E D V L S D G W D L Y A S L A Q G M L E P P S S M A - - A A ---- L G D Y ---- E V L V P -- L M S Y

```



```

CBF1 214 -----
CBF2 217 -----
CBF3 217 -----
CBF4 223 DE---
CBF5 208 RY---
CBF6 182 -----
TaCBF14 219 -----
TaCBF11 220 -----
TaCBF1 220 -----
TaCBF10 220 -----
TaCBF4 212 --LFD
TaCBF5 267 --LFD
TaCBF7 271 SHLFN
TaCBF8 220 --LFD
TaCBF2 226 -----
TaCBF6 246 -----
TaCBF15 246 -----
TaCBF13 246 -----
TaCBF9 242 -----
TaCBF12 235 -----
TaCBF3 237 QS---

```

Figure 3

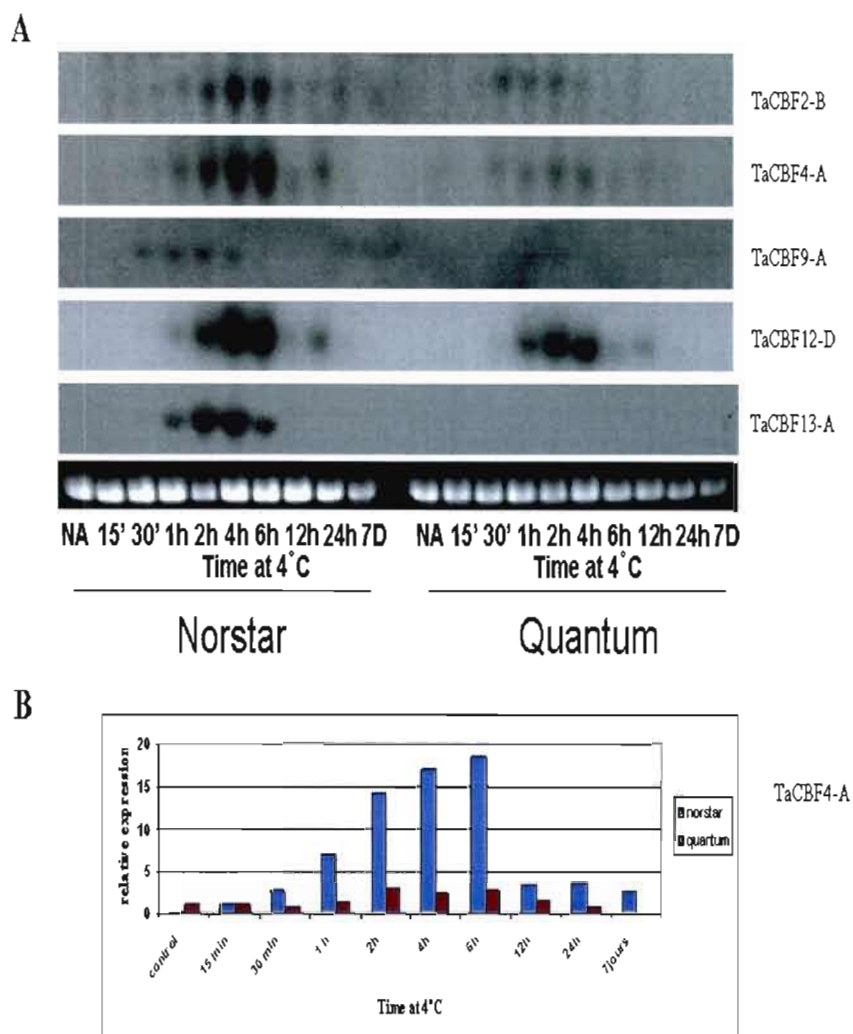


Figure 5

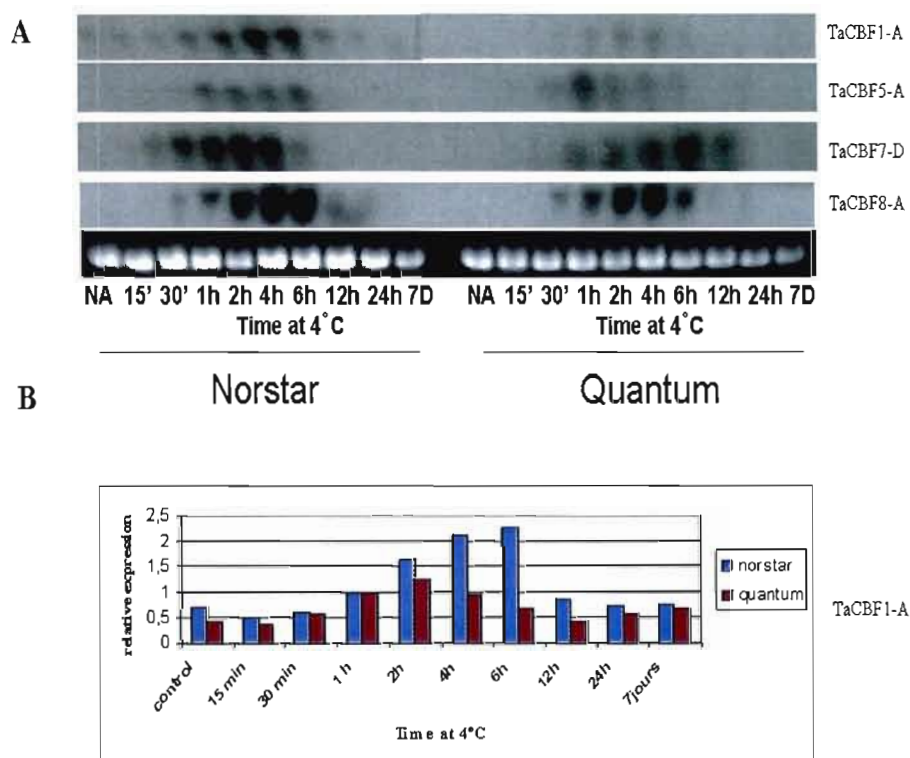


Figure 6

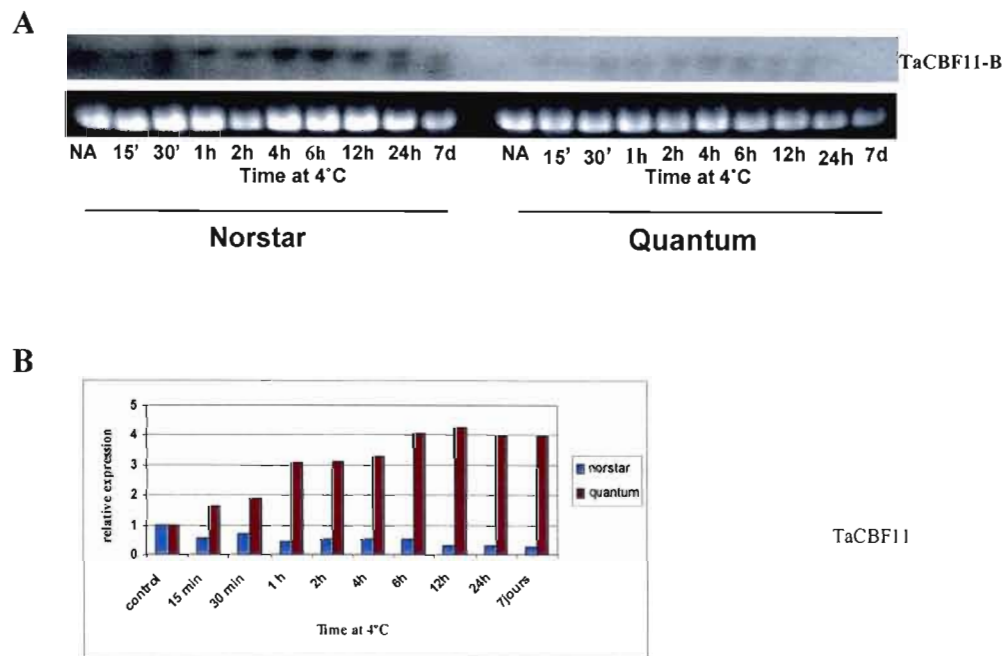


Figure 7

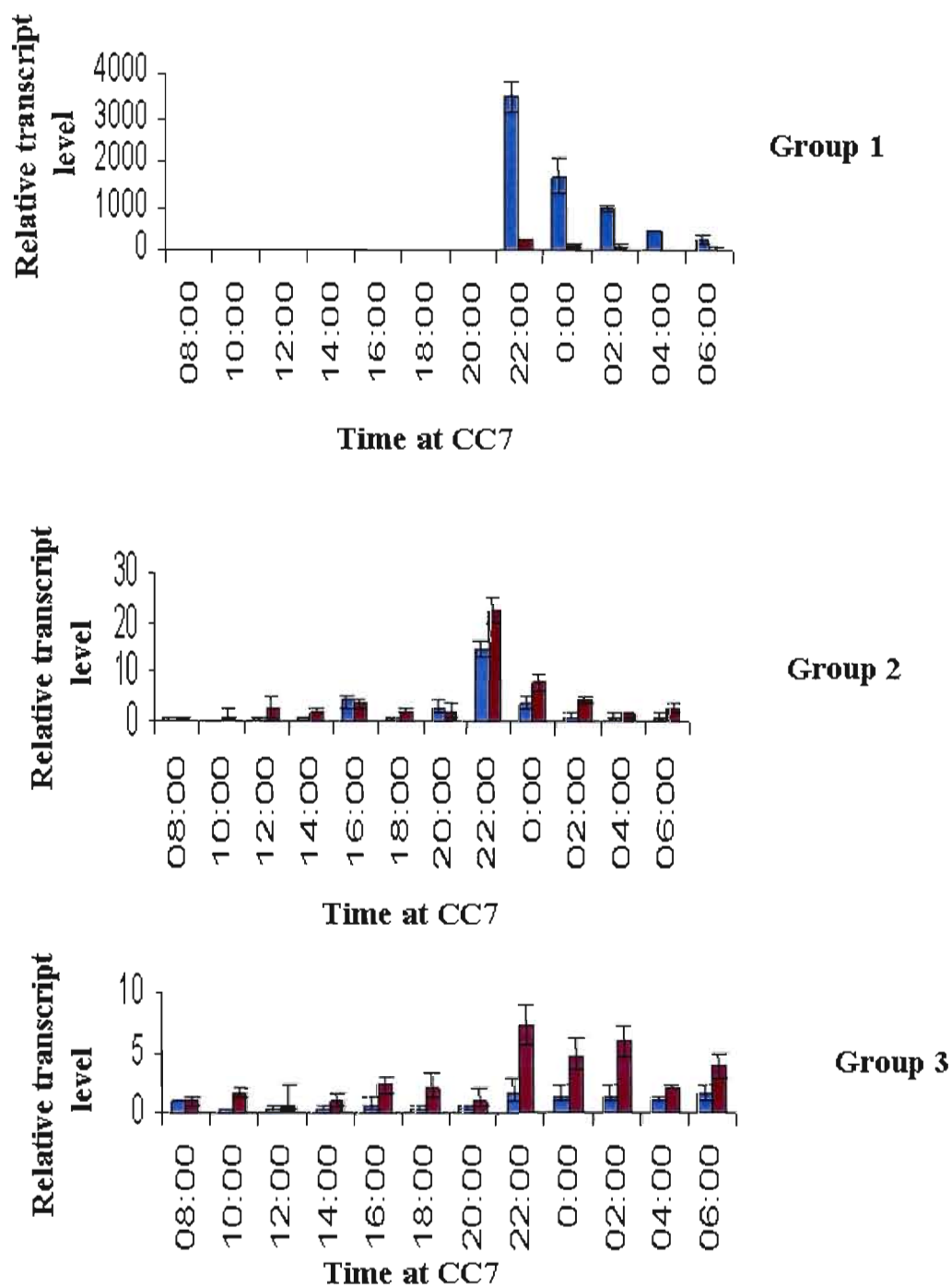


Figure 8

Chapitre II : Expression et purification DE LA PROTEINE *TaCBF1*

Introduction

Il a été démontré qu'un élément *cis* contenant un motif CCGAC appelé CRT (C-repeat) est impliqué dans l'expression de nombreux gènes induits au froid (Stockinger *et al.*, 1997). Ce même motif est aussi impliqué dans l'expression de gènes induits à la sécheresse. Des protéines liant CRT comme certains CBF chez *Arabidopsis* ont été rapporté comme activateurs de la transcription chez les promoteurs contenant le motif CCGAC.

Des études de liaisons ayant été effectuées avec les CBF1 et CBF2 de l'orge, des représentants des autres sous-groupes moins orthologues aux CBF1 et CBF2 seront utilisés pour éviter les redondances. Je vais produire et purifier la protéine de gène *TaCBF1* (première protéine a donnée des résultats intéressants). Ces vecteurs ont été choisie car les constructions dans les vecteurs pDEST peuvent être transformer dans différents systèmes (levures, cellules de plantes, bactéries..). D'autres kits comme TnT Quick coupled transcription translation (permet la transcription et la traduction *in vitro*) ou les vecteurs pET44 (fusion NusA Novagen) peuvent être utilisés pour produire des protéines recombinantes solubles. Ces dernières pourront être utilisées pour les études de liaison des CBF à un élément *cis* homologue à l'élément CRT.

Matériel et méthodes

J'ai sous cloné le cDNA *TaCBF1* dans un plasmide ENTRY vector (pENTR) pour faire la réaction de recombinaison des plasmides dans des vecteurs pDEST (vecteur d'expression contenant un tag His ou des stops, exemple pDEST14, 23, 41), en suivant les instructions du manufacturier pour la technologie Gateway. Il y a ensuite eu transformation des vecteurs d'expression, contenant la séquences des gènes, dans des bactéries compétentes (TOP10 ou BL21). Ces dernières ont été mises en culture et les protéines induites avec de l'arabinose. L'induction a été vérifiée grâce à un gel SDS-PAGE avec des bactéries non induites comme contrôle. Le gel de protéines a été utilisé pour confirmer que la bonne protéine a été induite grâce à une spectroscopie de masse. Les protéines ont été purifiées grâce à une colonne Ni-NTA qui correspond à la fusion utilisée (HIS-tag).

Après plusieurs essais avec la technologie Gateway, les protéines exprimées restaient insolubles et donc il aura fallu faire une purification en condition dénaturante. Nous avons donc changé de vecteur pour le système PET qui facilite la solubilisation de protéines insolubles, grâce à un tag NusA fusionné à la séquence. Ce tag code pour une protéine de 495aa qui est très soluble, ce qui permettrait d'augmenter la solubilité de protéines insolubles.

L'ADNc du *TaCBF1* a été sous cloné dans le vecteur pET-44 dirigé par un promoteur T7. Ce vecteur contient un His tag en N et C-terminal et le vecteur contient aussi pour un tag NusA interne.

Des PCR ont été effectués pour introduire des sites de restriction dans la séquence de la protéine. Les amorces utilisées pour l'amplification ainsi que les

sites de restriction introduits se trouvent ci-dessous.

	Séquence	enzyme
CBF1-9F1	5'ATGAGCTCATGGACACCGCCGCTGCCGGC3'	SacI
CBF1-9R1 stop	5'ATGGTACCTTAGTTCCAAAGCGGCGTGTA3'	KpnI
CBF1-9R2	5'ATGGTACCTTTGTTCCAAAGCGGCGTGTA3'	KpnI
COLIDOWN-R1	5'TTCACTTCTGAGTTCGGCATGGGGT3'	
S-tag-F1	5'GAACGCCAGCACATGGACTCGC3'	

Les fragments amplifiés (CBF1-9F1 et CBF1-9R1 stop) ont ensuite été clonés dans le vecteur pDRIVE puis digérés et sous-clonés dans le vecteur pET44b.

Les vecteurs d'expression, contenant l'ADNc, ont ensuite été transformés dans des bactéries compétentes Rosetta qui facilitent la formation de pont disulfures dans le cytoplasme facilitant ainsi le bon repliement de la protéine exprimée. Après avoir confirmé par séquençage la séquence des constructions, les bactéries ont été mises en culture dans 3 ml de LB-Amp et incubées o/n à 37°C avec agitation. Le lendemain, 2ml de la pré-culture a été inoculée dans 200 ml de LB-Amp-Cm-Kan puis incubés à 37°C pendant 3h. Une DO_{600} atteinte de 0.5-0.6 (phase log de croissance), 2 ml d'IPTG 100mM a été ajouté (concentration finale 1 mM) puis l'incubation a été poursuivie pendant une période optimale de 6 heures à 30°C. La culture a ensuite été centrifugée à V_{max} (10 500 rpm) pendant 10 min à 4°C. Le surnageant a été jeté et les culots congelés.

Les culots ont été ensuite décongelés puis suspendus dans 1 ml de tampon de lyse (350 mM NaH₂PO₄, 300mM NaCl, 20mM Imidazole pH 8.0) et 10 µl de lysosyme par ml de suspension (concentration finale 1 mg/ml) puis incubé pendant 30 min sur glace. Selon les besoins, 1 autre ml de tampon de lyse peut être rajouté afin de récupérer un maximum de bactéries. De la benzamidine et du PMSF aux concentrations de 1 mM ont été ajouté avant la sonication.

La suspension a été soniqué sur glace 6 fois 10 secondes avec des périodes de refroidissement de 10 secondes. Le lysat a été centrifugé à V_{max} pendant 5 min à 4°C puis le surnageant récupéré dans un tube de culture 13 ml propre et le culot gardé à 4°C (fraction culot). 1 ml de matrice Ni-NTA 50% a été ajouté à 4 ml de surnageant et incubé avec agitation douce (200 rpm) à 4°C pendant 1h.

Une fois que la résine a été transférer dans une colonne Bio-Rad et centrifuger à 1000 x g pendant 2 min à 4°C, le surnageant a été prélevé délicatement dans un tube propre et gardé à 4°C (fraction surnageant de lyse).

La résine a ensuite été suspendue doucement dans 2 ml de Wash Buffer (50 mM NaH₂PO₄, 300mM NaCl, 20mM Imidazole pH 8.0) puis agité pendant 10 min et centrifugé à 1000 x g pendant 2 min à 4°C, et le surnageant récupéré dans un tube propre et gardé à 4°C (fraction Wash 1). Cette manipulation a été répétée 2 fois ce qui nous a donné fractions Wash 2 et 3.

Les protéines ont été élué 4 fois avec 0.5 ml de tampon d'éluion (50 mM NaH₂PO₄, 300mM NaCl, 250mM Imidazole pH 8.0), puis agité pendant 15 minutes à 200 rpm à 4°C et centrifugé 1000 x g 2 min. Les fractions ont été collectées dans 4 tubes différents (Fraction Éluion 1, 2, 3 et 4). La quantité de protéines produites a été déterminée avec la trousse Rc-Dc (Bio-Rad).

Une électrophorèse sur gel de polyacrylamide a ensuite été réalisée pour vérifier les différentes fractions récupérées pendant la production et purification de protéines. Une partie des différentes fractions ont été mélangées avec 20-60 μ l de tampon de Laemmli 2-3X, chauffé à 95°C pendant 5 min puis déposés sur un gel de 12% acrylamide.

Un immunobuvardage (Houde et al, 1992) a ensuite été effectué. Après le transfert du gel sur membrane et saturation de cette dernière avec du blotto (1X PBS, 0.2% tween-20, 4% lait en poudre), elle a été incubée avec l'anticorps anti-CBF (dilution 1/50000) préparée dans une solution fraîche de blotto pendant une heure avec agitation. La membrane a ensuite été lavée 3 fois pendant 10 min dans du 1X PBS, 0.2% tween-20. Le deuxième anticorps (anti-IgG couplé à la peroxydase) a été ajouté puis un dernier lavage a été effectué. Les protéines détectées avec l'anticorps ont été révélées grâce au kit ECL detection kit (Amersham Bioscience) et visualisées sur film Kodak XRP-5.

Résultats

Pour vérifier si l'ADNc de *TaCBF1* code bien pour une protéine AP2, la protéine a été exprimée grâce à un vecteur pET. La bande de la protéine CBF1 observée sur gel de protéine, partiellement purifiée, sortait à environ 92-97kDA (tag 66kDA + CBF 25kDA) (figure 1a). Le immunobuvardage fait avec l'anticorps anti-CBF a confirmé la production d'une protéine CBF (figure 1b) malgré l'existence d'un signal non spécifique de plus faible poids moléculaire. Le dosage de la protéine exprimée a révélé une concentration approximative de 3mg/ml.

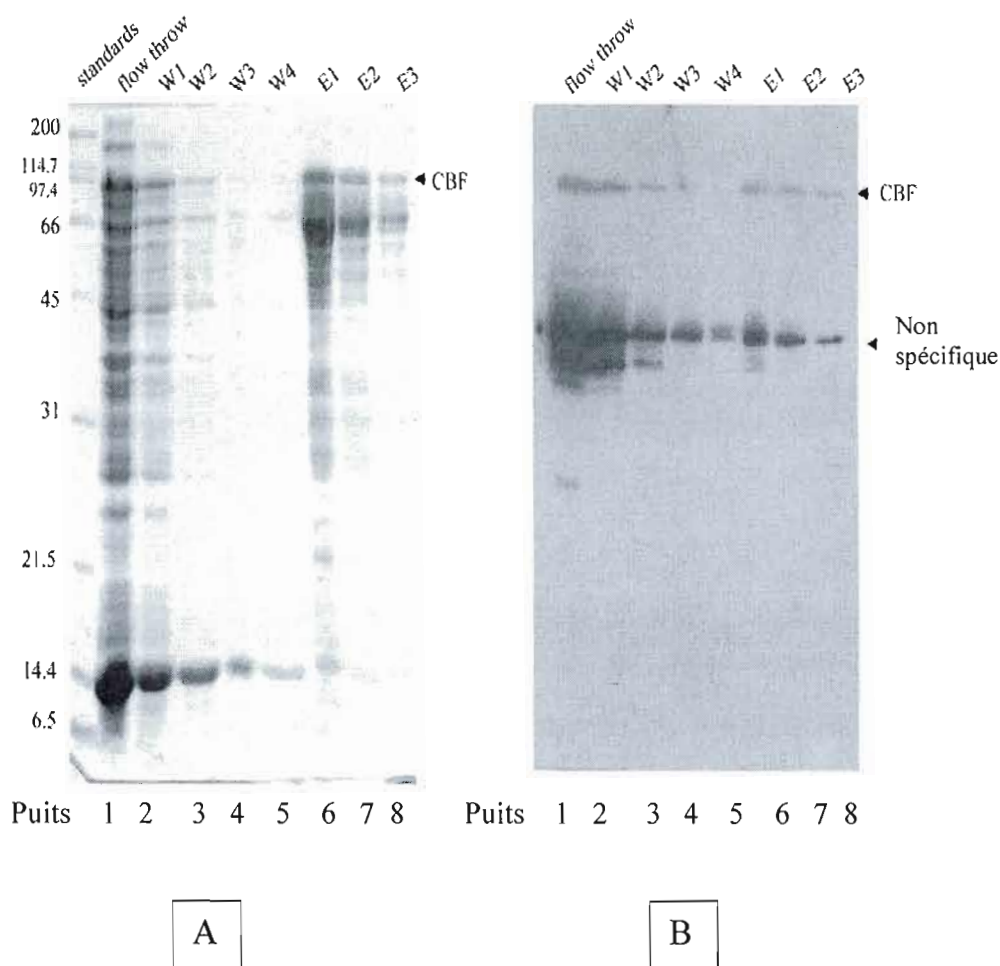


Figure 1: Expression de la protéine recombinante de *TaCBF1* dans *E.coli*

A. Séparation sur gel SDS-PAGE des fractions éluées. L'DNAc de *TaCBF1* a été inséré dans le vecteur pET44b et transformé par électroporation dans *E.coli*. Après induction avec IPTG, les bactéries ont poussé pendant 6 heures puis ont été récoltées. Le culot a été solubilisé et soniqué. Les protéines ont été purifiées grâce à

une colonne et séparées sur un gel 12% SDS-PAGE. Les protéines ont été détectées avec du Bleu de Coomassie. La flèche indique la position de la fusion NusA :: CBF.

B. L'analyse des fractions éluées a été faite grâce à un western blot. Les protéines ont été transférées sur une membrane PVDF et détectées avec l'anticorps anti-CBF.

Puit 1: flow throw. Puit 2: lavage fraction w1. Puit 3 : lavage fraction w2. Puit 4 : lavage fraction w3. Puit 5: lavage fraction w4. Puit 6 : élution fraction E1. Puit 7 : élution fraction E2. Puit 8 : élution fraction E3. La première flèche indique les protéines fusion NusA :: CBF purifiées et la deuxième flèche les interactions non spécifiques de l'anticorps.

Conclusion

Ce travail présente l'isolation et la caractérisation moléculaire d'une famille de gènes nommée *TaCBF* chez le blé, et appartenant à la famille des facteurs de transcription AP2. L'utilisation d'une approche génomique basée sur le séquençage à partir des banques d'ADNc de blé nous a permis d'identifier 37 cDNA contenant un domaine AP2. L'analyse des séquences a permis de les classer en 15 groupes appelés *TaCBF1-15* composant la famille CBF chez le blé. Chez *Arabidopsis*, les membres de cette famille ont des fonctions essentielles dans le développement de la tolérance aux stress abiotiques. Cela laisse supposer que les *TaCBFs* chez le blé auraient un rôle similaire.

Des études de cartographie chromosomique utilisant des lignées de délétion et de substitutions ont permis de positionner les *TaCBFs* sur les groupes de chromosomes 2 et 5, ce dernier étant reconnu pour être riche en gènes régulateurs de processus d'adaptation et de développement de la tolérance au gel chez le blé. Le positionnement des *TaCBFs* sur le chromosome 5 ainsi que la forte homologie des *TaCBFs* avec les *ArCBF1-6* nous a permis de prédire que les protéines pourraient être associées à la tolérance au froid chez le blé, ceci en déclenchant l'expression de nombreux gènes COR qui ont un rôle clé dans la protection de la plante contre le gel.

Afin de corroborer cette hypothèse, nous avons cherché à savoir si les messagers des *TaCBFs* étaient induits chez les plants de blé acclimatés. Les analyses des patrons d'expression ont révélé une accumulation rapide débutant 30 min après l'exposition au froid, pour atteindre un maximum après 4 à 6 heures suivie d'une baisse progressive de l'expression. Cette dernière était le plus souvent supérieure chez les cultivars d'hiver que chez les cultivars de printemps. Ceci permet d'associer l'amplitude de l'expression avec les plantes tolérantes au froid. Finalement, les analyses d'expression ont permis de déterminer que les messagers s'accumulent

exclusivement aux basses températures ; les autres stress abiotiques, soit la salinité, le choc thermique et la déshydratation n'induisent pas de façon significative l'expression des *TaCBF*. L'accumulation de *TaCBF* chez plusieurs génotypes de blé acclimatés au froid est fortement associée avec les gènes COR induits et de ce fait à la capacité de développer la tolérance au gel chez ces espèces végétales.

A la vue de ces résultats, il devient nécessaire de mieux caractériser cette famille de facteurs de transcription CBF chez les céréales. Avec différents projets de séquençage de génome de plantes et de génération de banques d'EST (Expressed Sequences Tags), de nombreux membres de cette famille ont été identifiés et clonés, mais très peu chez le blé. La suite logique de ce projet consistera à approfondir la caractérisation de ces gènes et de faire une étude comparative avec d'autres gènes pour mieux comprendre l'évolution de ces gènes ainsi que leurs fonctions chez les eucaryotes.

L'isolement des promoteurs des *TaCBFs* a été entrepris ainsi que la caractérisation de leurs séquences. Leur activité transitoire va être testée pour identifier les éléments *cis* nécessaires à leur régulation. Des plantes transgéniques sont en cours pour tester la surexpression des gènes et vérifier leurs fonctions. L'identification des principales composantes de la tolérance au gel et de l'acclimatation au froid ainsi que la connaissance de leur rôle dans ce processus pourront être utilisés dans l'amélioration de la tolérance chez les plantes d'importance économique.

ANNEXES

TaCBF1 clone 9

```

1      atcccaactcgaccgctcaaaccaacctgcaactctcaacgcagcgactttccactagttt
1      S Q L D R S N Q P A T L N A A T F H - F
62     ttgacgctgcaactgaggacaccgcgctgccggctccccgcgtgaggggcacaggacg
21     L T L Q L M D T A A A G S P R E G H R T
122    gtgtgctcggagccgcccaagaggccggcaggcgacccaagttcaggagacgcgccac
41     V C S E P F K R P A G R T K F R E T R H
182    ccgctgtaccgcggcgctgcggcgccggggccggctcgggcagtggggtgtgcgaggttcgc
61     F L Y R G V R R R G R L G Q W V C E V R
242    gtgcgcggcgcaaggtacaggtctgtgctcggcaccttcaccactgccgagatggcg
81     V R G A Q G Y R L W L G T F T T A E M A
302    gcgcgcgcgcacgactccgcgctgctcgcgctcctcgaccgcgcgcctgcctcaacttc
101    A R A H D S A V L A L L D R A A C L N F
362    gccgactccgctggcggtgctgccgctcctcgcggtggctcgtcccgcttcagcagc
121    A D S A W R M L P V L A A G S S R F S S
422    gcgcgggagatcaaggacgcgcgtcccatcgccgtcctggagttccagcggcagcgccc
141    A R E I K D A V A I A V L E F Q R Q R P
482    gtcgtgtcaacgtcggagatgcacgacggcgaaaaggacgcccaaggctcgcgcgacgcg
161    V V S T S E M H D G E K D A Q G S P T P
542    agcgagctgtccacgtccagcgacttgttgacgagcactggtttggcggcatggacgcc
181    S E L S T S S D L L D E H W F G G M D A
602    ggctcgtactacgcgagcttggcgcaggggatgctcatggagccgcgtccgccagaacg
201    G S Y Y A S L A Q G M L M E P P S A R T
662    tggagcgaggatggcggcgaatacagcgccgtctacacgccgctttggaacaaattatcc
221    W S E D G G E Y S A V Y T P L W N - L S
722    gactaattaagccatgtacagttttggaaactactccctcggtaaactaatataagagca
241    D - L S H V Q F W K L L P R - T N I R A
782    tttaaatacattaaaatagtgatctaaacactcttatattagtttacggagggagtaggct
261    F K S L K - - S K H S Y I S L R R E - A
842    actagtgtgtgtgttcttccaaataggggaagatacagagtaagcatcaggagcaat
281    T S G C V L P N M G R Y R V S I R S N F
902    tttcccgatatttgaaagcttcgggtttctatttgcaggaaggactgaattgtctatca
301    F P V F E S F R F L F A G K D - I C L S
962    agaccaaagaaaaatgacc
321    R P K K N D

```

Annexes 1a

TaCBF2 D992

```

1      atacagccacctcggttacaccacaaaccactctcaacgccagctgcgaccgctggacacc
1      I Q P P R Y T T N H S Q R Q L R P M D T
61     aacgccgctggccgcagtttgacgggcaagagtacaggacggtgtggccggaggagcag
21     N A A W P Q F D G Q E Y R T V W P E E Q
121    gagtaccggacggtgtggtcggagccgccgaagcggcgccggcgccggaacaagtgcag
41     E Y R T V W S E F P K R R A G R N K L Q
181    gagacacgccaccagtggtaccgcggcgtgcgccgcgtggccgggaagggcagtggtg
61     E T R H E V Y R G V R R R G R E G Q W V
241    tgcgagctgcgctgcccggccgaagccggagttactccaggatctggcttggcaccttc
81     C E L R V P A G S R S Y S R I W L G T F
301    gccagtgcccagatggcggcgcgcgcgacgactcgccgcgctcgcgctctccggccgc
101    A S A Q M A A R A H D S A A L A L S G R
361    gacgcgtgcctcaacttcgcgcactccgcctggcgatgatgccgtccacgcagccggg
121    D A C L N F A D S A W R M M P V H A A G
421    tcgttcaagttggccgcgcgagagatcaaggacgccgtcgccgtggccctcaaggag
141    S F K L A A A Q E I K D A V A V A L K E
481    ttccaggagcagcagccctgccgacgagtcacggcgccgtcggtccacggccgaggag
161    F Q E Q Q R P A D E S T A P S S T A E E
541    agcgcgctctccatcatccccagcgacctgtcggggctcgacaatgagcactggatcggc
181    S A L S I I P S D L S G L D N E H W I G
601    ggcagtgaggccgggtcgtactacgcgagcttggcgcaggggatgctcatggagccgccg
201    G M E A G S Y Y A S L A Q G M L M E P P
661    gccgacggagcttggcgggaggaccgcgaacacgacgacggattcgacacgtcgctgtgg
221    A D G A W R E D R E H D D G F D T S L W
721    agctactagtgtgatcaactgattaagcaatgtaaagatctagagagtactgtagtgct
241    S Y - C D Q L I K Q C K D L E S T A S A
781    agattgtgtttcaccaaatatgggaagaagagagagcaagcatcgggggaaaggtttccc
261    R L C F T K Y G K K R E - A S G E R F P
841    cccaatgtgaaagcgctcgggttctactccccgaagggcacaatgagcttctttcttta
281    P N V K A L G F Y S R K G T N E L L S L
901    tttaataaataaatagagaaatgagcagtagtagatc
301    F N K - I E K - A V V D

```

Annexe 1b

TaCBF3 D917

```

1      ccaaaaccacttcacactcgagcaaacaccccaagagcaccaggacctccaagccggccaa
1      Q N H F T L E Q H P K S T Q D L Q A G Q
62     gaaacagagcatccgagcaaaacctctctgaattcaaggaagaagcgtgtccgatcaag
21     E T E H P S K T S L N S R K K M C P I K
122    agggagatgagcggggagtcggctcgccgtgcagcggggagagcttctactcgccctcc
41     R E M S G E S G S P C S G E S F Y S P S
182    acgtcgccggagaaccagcaggcgaggcaggcgcggtggacgtcgcgccggcgaagcgg
61     T S P E N Q Q A R Q A A W T S A P A K F
242    ccggcggggcggaccaagttcagggagacgcggcaccgggtgtaccgcggcggtgcggcgc
81     P A G R T K F R E T R H F V Y R G V R R
302    aggggcaatgccgggcgggtgggtgtgctgaggtgcgcgtgcccggcaggcgcgggagcagg
101    R G N A G R W V C E V R V P G R R G S R
362    ctctggtcggcaccttcgacacgcggcaggccgcgcgcgcgccaacgacgcggccatg
121    L W L G T F D T A E A A A R A N D A A M
422    atcgcgctctccgcggggggcgcggtgctcacttcgcgcgactcgccgcgagctgctc
141    I A L S A G G A G C L N F A D S A E L L
482    gccgtgcggcagcctcctcctaccgcagcctcgacgaggtccgccacgcgcgtcggtggag
161    A V P A A S S Y R S L D E V R H A V V E
542    gccgtcgaggacttgctgcggcgcgaggcgatcgccgaggacgacgcgctctccggcacc
181    A V E D L L R R E A I A E D D A L S G T
602    tctcgtcgcgccctcctccctcaccgacgacgggtcgctcctcctcgccgctgcccag
201    S S S A P S S L T D D G S S S S P L P E
662    gaggactcgccgttcgagctggacgtgctgagcgacatgggctgggacctgtactacgcg
221    E D S P F E L D V L S D M G W D L Y Y A
722    agcctggcgcgaggcgatgctcatggcgccgccttcttccatggccgcggcgctcgccgat
241    S L A Q A M L M A P P S S M A A A L G D
782    tacggcgaggctgatgtgccactctggagctaccagagcagcaatagttcgcgccagtt
261    Y G E V D V P L W S Y Q S - Q - F A P V
842    agaataattttaccttttctttctgtccgtcgtcttggcttccgatcgatgccaaaatttt
281    R I F Y L S F L S V V L A S D R C Q N F
902    ggtgctgctttagggtcaccgcttgcagtttctagtaatgtgatggtgcgaattgggg
301    G A A L G S P L A V S S N V M V R I G

```

Annexe 1c

TaCBF4 D990

```

1      cgctcctctctcaagtgtctcaacgcagcagctaaacacgctaagctacctcgctgctta
1      R S S L K C L N A A A K H A K L P R C L
61     attactccacagtcgacaggctcccggcgacactgcgatcggttcgacgacgcccgcgac
21     I T P Q S T G S R R H C D R S M D A A D
121    gccgcctccccgtgtgatggccacaggacgggtgtggtcggagccgcgaagcggccggcc
41     A A S P C D G H R T V W S E P F K R P A
181    ggccggaccaagttcaaggagacgcgccaccgcgtgtaccgcgcggtgcggcgacggggc
61     G P T K F K E T R H F L Y R G V R R R G
241    cccgcccggcgggtgggtgtgcgaggtgcgcgtgctcgggatgaggggctccaggctctgg
81     P A G R W V C E V R V L G M R G S R L W
301    ctcggcaccttcaccaccgctgagatggcagcgcgcgcgacgacgcccggcttctcgcg
101    L G T F T T A E M A A R A H D A A V L A
361    ctctctgggcgcgcgcttgtctcaacttcgccgactccgcctggcggtatgctccccgtc
121    L S G R A A C L N F A D S A W R M L P V
421    ctgcgcggcccgttcagcaccgccaaggagatcaaggacgcgctcgccgtcgccgtcctg
141    L A G P F S T A K E I K D A V A V A V L
481    gcgttccaaaggcagcaccgggtcgcgccatggcaccattgtcccctgcgcggacaacc
161    A F Q R Q H R V A S M A P L S P A R T T
541    gatgacgagaaggaaatcgatggctcgccggcgccgagcgccctgttcatgtccagcgag
181    D D E K E I D G S P A P S A L F M S S E
601    ctggtgaatgagcactggtttggcgcatggtatgccgqatcggttctactcggagggcctg
201    L L N E H W F G G M D A G S F Y S E G L
661    ttcattggagtcgcccgcacaccagaccgtggcggaagacctcgagctctgtggcgctcga
221    F M E S P D T R P W R E D L E L C G V E
721    acaccgcatggagctacttgttcgacaaagcagttaagttaaatgtttgtatagatag
241    T P P W S Y L F D - S S - V K C L Y R -
781    tttctctcttctgttccaccaaataatgggaggaacagagggggatattttcccatatg
261    F L S S V P P N M G G N R G G Y F P H M
841    tccgtatgtgctcggttcattgtg
281    S V C A R S L

```

Annexe 1d

TaCBF5 D921

```

1      tccaccgatcgatcaaaacctctcaacacagccgctgattcttccagtactcctgtccac
1      P P I D Q N L S T Q P L I L P V L L L H
62     acctccacgagcgtctccgccagctctcgactcagatgacgtcgccgacatcgctcc
21     T S H E R L R Q L S T Q M D V A D I A S
122    ccgtctggccagcaggggcatcgggacaccggacggtgtcgtcggagccgccaagcgc
41     P S G Q Q G H R G H R T V S S E P F K F
182    cccgcggggcgaccaaagttccacgagacgcgccaccgcgtgtaccgcggcgtgcggcgc
61     P A G R T K F H E T R H F L Y R G V R R
242    cgtggccgcgtcgggcagtggtgtgcgaggtgcgcgtgcctgggatcaagggtccagg
81     R G R V G Q W V C E V R V P G I K G S R
302    ctctggctcggcaccttcaacacggccgagatggcggcgcgcgcgacgacgcccgtgtg
101    L W L G T F N T A E M A A R A H D A A V
362    ctgcgctcatcgccgcgcgcctgcctcaacttcgccgactccgcatggcgcatgtg
121    L A L I G R A A C L N F A D S A W R M L
422    cccgtgctcgcggccgggtcgttcggtttcggcagcgcgagcgagatcaaggccgcgtc
141    P V L A A G S F G F G S A S E I K A A V
482    gccgtggccgtcgtcgcgttccagcgaagcagattattccggtcgcgcgtcgcgtcgtg
161    A V A V V A F Q R K Q I I P V A V A V V
542    gcgctccagcagcagcaggttcgggtcgcgtcgcgtcgtggcgctccagcagaagcag
181    A L Q Q Q Q V P V A V A V V A L Q Q K Q
602    gttccggtcgcgtggccgctcgtggcgctccagcagctgcaggttcgggtcgcgtggcc
201    V P V A V A V V A L Q Q L Q V P V A V A
662    gtcgtggcgctccagcagcagcagattattcttccagtgcgtgcctggcgccgagttt
221    V V A L Q Q Q Q I I L P V A C L A P E F
722    tacatgtcttccggcgacctgttgagcttgacgaggagcagtggtttggcgcatggac
241    Y M S S G D L L E L D E E Q W F G G M D
782    gctgggtcgtactacgcgagcttggcgcaggggatgctcgtggcgccgacgaaaga
261    A G S Y Y A S L A Q G M L V A P P D E R
842    gcgaggccggagcacggcgagcagagcgcgtccagacgcccgtatggagctgcttgc
281    A R P E H G E Q S G V Q T P L W S C L L
902    gactaattcagaaactactttcaagttgtagatttaggaagaacatactaggcagtttgt
301    D - F R T T F K L - I - E E T Y - A V C
962    tccttcattttagtagtagtatttggtagttcccggtttttacttttggggagaagggcta
321    S F I C S S I C G V P G F Y F W G E G L
1202   gattgcttaatacaccaattattggtttg
341    D C L I H Q L L V L

```

Annexe 1e

TaCBF6 L3C105_F16-9

```

1      cagctagctcacactcctcactgagctaagcaccaacctcgactgccaaggaagccacc
1      Q L A H T P H - A K H Q P R L P K E A T
61     tgccaatcacccagcactctgccggtagccatggacatgggccttgaggctcgcagctcc
21     C Q S P S T L P V A M D M G L E V S S S
121    tccccatcctcctcgctcgctggccaagcgccccgcggggcgccaagtccgcgagacg
41     S P S S S S L A K R P A G R T K F R E T
181    cggcaccctgtgtatcgtggcgtgcggcgccggggcaacgccaacgggtgggtctgcgag
61     E H E V Y R G V R R R G N A Q R W V C E
241    gtgcgcgtccctggcaagcgcggtcgctcggtctggctcgggacttacgccacggccgag
81     V R V P G K R G A R L W L G T Y A T A E
301    attgcagcgcgcgccaacgacgcgcgccatgctcgccctgggcggcgctccgcgcgctc
101    I A A R A N D A A M L A L G G R S A A L
361    ctcaacttccccgactccgcgtggctgctcgctgtgccgtccgcgcactccgatctcgcc
121    L N F P D S A W L L A V P S A H S D L A
421    gacgtccggcgcgcggtcgagccgtcgcgatttgacgcgacgggaggtcgccggt
141    D V R R A A V E A V A D L Q R R E A A G
481    ggggtccatcacgccaccgccaccgccaccgcccgaggaggcctcctgtggcgctcct
161    G S I T A T A T A T A A E E A S C G A P
541    gcagaatcgctcgtcgtgagctgacgatgccggttcgctcggagacgtcgaaacctccgcc
181    A E S S S E S D D A G S S E T S K P S A
601    gatggagacttcgcggtgccggcggaatggacatcgaaatgttcagtaggcttgacttg
201    D G D F A V P G G M D I E M F S R L D L
661    ttccccgaaatggacttgggtcgtactacgcgagcctcgcgaggcgctgctcatggac
221    F P E M D L G S Y Y A S L A E A L L M D
721    ccgctccgggtggcgactggcaccggcgcgactgggataacggagagtgccggcgaggcc
241    P P P V A T G T G A Y W D N G E C G E A
781    gagggggcaactgagttcgcgctctggagcagtagccgattctgcttttgttttgactc
261    E G A T E F A L W S - - P I L L L F - L
841    tgtagcgtctttttccctttgcaaagttttccaatttcgttttt
281    C S V F F P F A K F S N F V F

```

Annexe 1f

TaCBF7 L6B020-O14F1-L1-68

```

1      ctcagttctcaccaccccaaaacacgcccgtgaaacctccagttctcagcacacatcgac
1      L S S H P P K H R R - N L Q F S A H I D
61     ctgctcgaccgtctccactagcgctcgccaagctggacgtcgccgacgcccgcctccccg
21     L L D R L H - R S S K M D V A D A A S P
121    tctggccaggagcagcaggggcacaggacgggtgctgctgagccgccaagcgacccgcg
41     S G Q E Q Q G H R T V S S E P F K R P A
181    gggcggaccaagggtacatgagacgcgcacccgctgtaccgcggtgcggcagcggggc
61     E R T K V H E T R H E L Y R G V R Q R G
241    cgggtcgggcagtggtgtgctgaggtgctgctggcgggggtgaagggtccagggtctggtg
81     R V G Q W V C E V R V A G V K G S R L W
301    ctcgccaccttcaccaccgcccagatggcgcgcgcgcgacgacgcccgtgctcgcg
101    L G T F T T A E M A A R A H D A A V L A
361    ctctccggccgggctgctgctcaacttcgcccactccgctggcgcatgctgcccgtg
121    L S G R A A C L N F A D S A W R M L P V
421    ctcgcgccgggtctttcggttcggcagcgcgcgagatcaaggccgcccgtcgccgtc
141    L A A G S F G F G S A R E I K A A V A V
481    gccgtcggtgctttccagaaggagcagattatcccagtcgcccgtcgccgtggtggcgatc
161    A V V A F Q K E Q I I P V A V A V V A I
541    cagaagcagcagattatcccagtcgctgctgcccgtcggtggcgatccagaagcagcagatt
181    Q K Q Q I I P V A V A V V A I Q K Q Q I
601    atcccgtcgccgtcggtggcgctccagaagcagcagattccagtgccgctcgccctcggtgg
201    I P S P S W R S R S S R F Q W P S P S W
661    cgctccaggaacagcaggtcccagtcgcccgtcgccgtcggtggcgcttcacgagcagcag
221    R S R N S R S Q S P S P S W R F I G S R
721    ttccagtcgctgcccgggaacctccggcccgggcagcgctctgttttacatgctcgcca
241    F Q S R A R E P P A R A A L C F T C R P
781    gcgacttggtggagctcgacgaggagcagtggttggcgcatggaggccgggtcggtact
261    A T C W S S T R S S G L A A W R P G R T
841    acgcgagcttgccgcaggggatgctcggtggcaccgcccggacgaaagagcgaggacggagg
281    T R A W R R G C S W H R R T K E R G R R
901    acggcgagcagagcggcgctccagacgcccgtatggagccacttggtcacctaatactagca
301    T A S R A A S R R R Y G A T C S P N L A
961    gtgaaagttgtactccctctgtccggaattacttggtgccgaaatggattaaaatgaatg
321    V - S C T P S V R N Y L C R N G L K - M
1021  tatttagaactaaaatacgtctagatacatcaatttctc
341  Y L E L K Y V - I H Q F L

```

Annexe 1g

TaCBF8 L6B012-G16-13

```

1      agtaaaaacaccatcgctcaacctgctcaaccatctccactagctctcgactcagatggac
1      V K T P S L N L L N H L H - L S T Q M D
62     gtcgccgacgctgctccaagtccggccaccaggagcagggtcacaggacgggtgcatcg
21     V A D A A S K S G H Q E Q G H R T V S S
122    gagccgcccgaagcggccggcgggcgccgaccaagttccgtgagacacgacaccgctgtac
41     E P F K R P A G R T K F R E T R H E L Y
182    cgcgcggtgcggcgccgtggcggggtcgggcagtggtgtgaggtgcggtgcgcagga
61     R G V R R R G R V G Q W V C E V R V P G
242    gtgaagggtccaggctctggctcggcaccttcaccaccgcccagatggcgcgcgcgcg
81     V K G S R L W L G T F T T A E M A A R A
302    cagcagcccggtgctcgcgctctccggcgccgcccgcctgctcaacttcgcccgactcc
101    H D A A V L A L S G R A A C L N F A D S
362    gcatggcggtatgctgcccgtgctcgcggccgggtcggttctgcttcggcagcgcgcgggag
121    A W R M L P V L A A G S F C F G S A R E
422    atcaagcttgccgtcgccgtcgccgtcgctcgcgttccagctgcagcagaatattcctcca
141    I K L A V A V A V V A F Q L Q Q N I P P
482    gcgcgtgtccaacggtggagccgacccagacagcccagcaactctctgttctacatgtcg
161    A A C P T V E P T D S P S N S L F Y M S
542    tccggcgacctgctggagcgcgacgaggagcactgggtggcgccatggacgccgggtcg
181    S G D L L E L D E E H W F G G M D A G S
602    tactacgacagcttgccgaggggatgctcgtggagccccggacgacagagcgaggcg
201    Y Y D S L A Q G M L V E P P D D R A R R
662    gaggacgccgagcagacggcgctcgagacaccgacgccgctatggagctattgtttgac
221    E D A E Q T G V E T P T P L W S Y L F D
722    Taaattcagcagcagtgtaaagttgtcgatagttctgttgttcttcccgatctgggaa
241    - F S T Q C K V V D S S V V F F P I W E
782    caaactgagtaggcagttttttttgaacgttgtagtaggcagtttgaagctcccggtt
261    Q T E - A V F F - T L - - A V - S S R F
842    ctactttgtgagaattggaccct
281    L L L - E L D P

```

Annexe 1h

TaCBF9 L6B002-G24-33

```

1      ataaccaacactcctcagctcaagctcagcctgctctcacactccatagatcgacctcggc
1      - P T L L S S S S A C S H T P - I D L G
62     ccaccgtacgtaagctcgagccagcaccgaccacctgcagtaccaatggacaccggc
21     P P Y V S S S Q H R P P A V P M D M T G
121    tccgaccagcagcggagctccccgtcctcgccgtcgctcctcctcgcacctgaagcgtcct
41     S D Q Q R S S P S S P S S S S H L K E F
181    gccgggcgacccaagttcaaggagacgcgccaccgggtgtaccgcggtgcggcgccgc
61     A G R T K F K E T R H F V Y R G V R R R
241    ggcagcgccggccgggtgggtgtgcgaggtgcgcgtccccggcaagcgccggcagcggtc
81     G S A G R W V C E V R V P G K R G E R L
301    tggctcgccacgcacctcactgccgagggcgccgcgcaagcgacgatgccgccatgctc
101    W L G T H L T A E A A A Q A H D A A M L
361    ggctgatcgcccccctccaccctgtctcaacttcgccgactccgcttggctcctcgcc
121    G L I G P S T P C L N F A D S A W L L A
421    gtgccgtccgcgtctcgcactttgcagacgtccggcgccggtctcagcgtgtcgcg
141    V P S A L S D F A D V R R A A L S A V A
481    gacttccaacgccgggagggccgagcgccgcccaccacgtccctggccgccaccgtc
161    D F Q R R E A A S G A A T T S L A A T V
541    cccgtcgatgacggaagctgtagccaatccgcgcagtcgtccatggaaaataccggctcg
181    P V D D G S C S Q S A Q S S M E N T G S
601    tcgtggacgtcgctcgctcgctccctaccttctggcgacggaatgtttgcgggtgccggcc
201    S W T S S S S S L P S G D G M F A V P A
661    acaactgggctgcaacatgttcgagctggatatgtccggggaaatggacctggacacatac
221    T L G C N M F E L D M S G E M D L D T Y
721    tatgcgtacttcgcggaggggttctcctcgcagccgcgcaaccgcggtcgccggcgcg
241    Y A Y F A E G L L L E P P Q P P V A G A
781    tgctgggacaccgaagcgagcgagcagacgcgcgctctcgagctacaaacactcca
261    C W D T E G S G A D A A L S S Y - N T P
841    gccagctgctttgactctatgcatggaacaaaccgaagctgttcctgtcagcataataa
281    A S C F D S M H G T N R S C S L S A - -
901    aattgtggaacttaacc
301    N C G T - T

```

Annexe 1i

TaCBF10 C133R1L1-4

```

1      agacaccagagcatctcatggcaaaatagaaaaaatgtaacaaaagtagcaccgtagtcga
1      D T R A S H G K I E K M - Q K - H R S R
62     tcgtacgtccaagcgagagctaccgctagctacctcgatgaagcaaacgcgcgcatatg
21     S Y V Q A R A T A S Y L D E A N A R H M
122    aagctggctactgctcgatattgtaagtgaagtgagaaaccagctctgcatgagagtgt
41     K L A T A R Y C K - E - E T S L P - E C
182    cccttcgttccacttccatttccaccgagcgtccacgtcgtggcgggtatccaaccgt
61     P F V P L P F P P E R P R R G G Y P T R
242    tgcgggtaacccctcaacctcttctcctgcatactatatactcacctgcttgattctccg
81     C G - P L N L F S C I L Y T H L L D S P
302    cttctagttctcatcggagccccaatcaatcgaacctcgccagtcgccacaacggcttaa
101    L L V L I G A P I N R T S P V A T T A -
362    cgcacgccacccccacagtgagacgcgacgcgcctccctgtcagaccagcacaggacg
121    R T P P P Q M D A D A A S L S D Q H R T
422    gtgtggacggagcgcgcgaagaggcggcgaggagatcaagtacaaggagacgcgccac
141    V W T E P F K R P A G R I K Y K E T R F
482    ccgctgtaccgcggcgtgcggcgtcggggcggtacgggcggtgggtgtgaggtacgc
161    F L Y R G V R R R G R Y G R W V C E V R
542    gtgcggcgagcaaagagacaaggctctggtcggcaccttccgcaccgcccagatggcg
181    V R G S K E T R L W L G T F R T A E M A
602    gcgcgagcgcacgactccgcctccctcgctctctccggaagcgcgcctgcctcaacttc
201    A R A H D S A S L A L S G S A A C L N F
662    gccgattccgcctggcgtatgctgccgctcctcgccgctgggtcttccagcttcagcagc
221    A D S A W R M L P V L A A G S S S F S S
722    gcgcgggagatcaaggacccgctgcgcgctgctgctggcgttccagcggcagcgaccc
241    A R E I K D A V A V A V A F Q R Q R P
782    gtcgcgtcgacggcgcgagcggagaaaggacgtccaaggctcgccgacgcgcgagcgtg
261    V A S T A D G E K D V Q G S P T P S E L
842    tccacgtccagcgcacttgctggacgagcactgggttggcggcacgaatgccggatcatc
281    S T S S D L L D E H W F G G T N A G S Y
902    tactcgccgggcatgttcagtgagtcgcgcgagcagcctgaaaaccacgagctcggtggc
301    Y S P G M F M E S P E Q P E N H E L G G
962    ggcgacgtcaagacgcgcgtatggaagtttatagttggttctatacttctgtcccacc
321    G D V K T P L W - S L - L V L Y F C P T
1022   agatatgggaagaacagaggggaatatttccctcatatggtcatgtgtgttcgttcac
341    R Y G K K Q R G I F S L I W S C V F V H
1082   attgaaagcctccgg
361    I E S L R

```

Annexe 1j

Annexe 1k

TaCBF12 L6B007_F24-19

```

1      acccaacacctcaaccacactctgcacacccacactcgtctacccaacacgtacctcaagc
1      T Q H L N H T L H T T L V Y P T R T S S
61     tcaagctcaagccagccataccgcgaaaaaaaaagagagaaaaagctcaagctcagcta
21     S S S S Q P Y R E K K R E E K A Q A Q L
121    gagtgaccacatcccccaagcgatcggcactccgacagctagcagggacatgggcatc
41     E - P H P P K R S A L R Q L A M D M G I
181    aacggctggatcagctcccttcctcatcgacgtccgggcacgagctcggggaggcggtg
61     N G W I S S P S S S T S G H E L G E A V
241    cctgtgtggtcgcaggcggcgaagcggcccgaggcgccacaaagttcaaggagacgcgc
81     P V W S Q A A K R P A G R T K F K E T R
301    caccgggtgtaccgtggcgtgcggcgccggggcagcgcgggcggtgggtgtgcgaggtg
101    H F V Y R G V R R R G S A G R W V C E V
361    cgcggtcccggaagcggcgagcggctctggtcgggacgtacgtcgccgagtc
121    R V P G K R G E R L W L G T Y V A A E S
421    gccgcgcgcgtcacgacgccgcgatgtcacccctgctcggacgtccccctccgcgcg
141    A A R A H D A A M L T L L G R S P S A A
481    gcgtgcctcaacttcccagactccgcgtggctgctcgtcatgccccgaggctctccgac
161    A C L N F P D S A W L L V M P P R L S D
541    ctggccgacgtccggcgcgcccatccaggccgtcgtgggttctctgcgctggaggcc
181    L A D V R R A A I Q A V V G F L R L E A
601    gccacggtggtcccgacgtcgacgagccacctccccgtgtacctgcccgtcgccgtg
201    A T V V P D V D E A T S P V Y L P S P V
661    gacaatgccgaccaagtgttccagggtgccaactttctccccgctgggcagcgacatgttc
221    D N A D Q V F Q V P T F S P L G S D M F
721    gagctcgacatgtctggggaaatggacctggacgcgtactacgcggttttcccagggg
241    E L D M S G E M D L D A Y Y A G F A Q G
781    atgctcctggagccgcccacgcgctcgtactgggagaaacggagaatgcggcgacggc
261    M L L E P P P T P S Y W E N G E C G D G
841    ggagcggccgcccgcctctggagctactgatgtctgtaccctgcagtgggttcacactga
281    G A A A G L W S Y - C L - P A V G S H -
901    ggattttgttatcccgattggaccaactgaaaactgctccaagcataataaaaatcggag
301    G F C Y P D W T N - K L L Q A - - K S E
961    tttgg
321    F

```

Annexe 11

Annexe 1m

TaCBF14 BQ744365R1L8-153

```

1      cacctcgaacggactgatcactgatggagtggcggtacagcggcgggccactcgtcgtcg
1      P R T D - S L M E W A Y S G G G H S S S
63     ggaacgaaatctccggcgccggcgccgggaggagggtcgtacacgacgggtgctcgtcg
21     G T K S P A A G G R E E G S Y M T V S S
123    gcgccgcccgaagcggcgggcgggaggaccaaggtcagggagacgaggcaccggtgtac
41     A P P K F R A G R T K V R E T R H T V Y
183    aagggggtgcgagcaggaaccccgccggtgggtctgagaggtgcgagcgccgacggg
61     K G V R S R N P G R W V C E V R E P H G
243    aagcagcggctgtggctcggcaccttcgacaccgcccagatggcgcgcgcgccgacgac
81     K Q R L W L G T F D T A E M A A R A H D
303    gtcgcccgcctcgcgctccgcggcgccgcgcgctgcctcaacttcgcggaactcgcgcgc
101    V A A L A L R G R A A C L N F A D S P R
363    acgctccgggtgcccgcgcagggggccgcccacgaggagatagccgcgcgcgcggtcgag
121    T L R V P P Q G A G H E E I R R A A V E
423    gcggccgagctgttccgcccggagcctgggcagcgcaatgcagctactaccgagggccg
141    A A E L F R P E P G Q R N A A T T E A P
483    gctgcttcaccggcagacgcggggaacgcggagctcgttgcaaactctccttaccacctc
161    A A S P A D A G N A E L V A N S P Y H L
543    atggatgggttagaatttcgaatgcagggtatcttgacatggcgacggcatgctgatc
181    M D G L E F E M Q G Y L D M A H G M L I
603    gagccaccgccaatggcggggcccgtcgacgtggatcgaggaggactacgactgcgaggtc
201    E P P P M A G P S T W I E E D Y D C E V
663    agcctgtggaactactgatgccgcgcacctgagccggctcgggcctggccttcatggctc
221    S L W N Y - C R A P E P A R A W P S W L
723    actgtcactgtacatacagctgaatcgcgcatccaggatgccctgtacttgtgggtgtg
241    T V T V H T A - Y R H P G C P V L V V C
783    aatgtttccctgctattctgacgaaaatgtcatggcaatgtgggtcaacaaacaattagt
261    N V S L L F - R K C H G N V G Q Q T I S
843    gcc
281    A

```

Annexe 1n

Table 1: Homology of sequence between TaCBF1 and the other TaCBF and ArCBF.

Genes	homology
TaCBF1	
TaCBF2	72%
TaCBF3	68%
TaCBF4	83%
TaCBF5	85%
TaCBF6	70%
TaCBF7	85%
TaCBF8	86%
TaCBF9	67%
TaCBF10	81%
TaCBF11	66%
TaCBF12	69%
TaCBF13	65%
TaCBF14	68%
TaCBF15	71%
ArCBF1	65%
ArCBF2	63%
ArCBF3	68%
ArCBF4	65%
ArCBF5	65%
ArCBF6	66%

Annexe 2

Table II. Primer sequences used in Northern blot experiments

Primer Name	Direction	Primer Sequence	Amplicon size (pb)
TaCBF1-A clone 9	Forward	5' TCCAGCGGAAGCAGATTATT 3'	237
TaCBF1-A clone 9	Reverse	5' CATTACCAAAGAGCGACTTGAACC 3'	
TaCBF2-A D992	Forward	5' TCGTACTACGCGAGCTTGGCGC 3'	267
TaCBF2-A D992	Reverse	5' TGTGCCCTTCCGGGAGTAGAAACC 3'	
TaCBF3-B D917	Forward	5' ACCGACGACGGGTCGTCCTCCT 3'	305
TaCBF3-B D917	Reverse	5' ACTGCAAGCGGTGACCCTAAAGCA 3'	
TaCBF4-A D990	Forward	5' GATGCCGGATCGTTCTACTC 3'	110
TaCBF4-A D990	Reverse	5' AAGTAGCTCCATGGCGGTGT 3'	
TaCBF5-A D921	Forward	5' GCGGAAGCAGATTATCCGGTCGC 3'	377
TaCBF5-A D921	Reverse	5' CGTCTGGACGCCGCTCTGCT 3'	
TaCBF6-A L5C016_A21-37	Forward	5' GAGGCGCTGCTCATGGAC 3'	150
TaCBF6-A L5C016_A21-37	Reverse	5' AAAGACGCTACAGAGTCAAAACAAA 3'	
TaCBF7-A L6B020-O14F1-L1-68	Forward	5' CTCAGTTCTACCCACCCAAACA 3'	210
TaCBF7-A L6B020-O14F1-L1-68	Reverse	5' GTGGCGCTCTCATGTACCTT 3'	210
TaCBF8-A L6B012-G16-13	Forward	5' TCGACGAGGAGCACTGGTTT 3'	304
TaCBF8-A L6B012-G16-13	Reverse	5' AGGGTCCAATTCTCACAAAAGTAG 3'	304
TaCBF9-A L6B002-G24-33	Forward	5' TACCTTCTGGCGACGGAATGTTT 3'	224
TaCBF9-A L6B002-G24-33	Reverse	5' GCAGCTGGCTGGAGTGTTT TAGTA 3'	224
TaCBF10-B C133R1L1-4	Forward	5' AGCTGTCCACGTCCAGCGATTT 3'	300

TaCBF10-B C133R1L1-4	Reverse	5 TCACAGAACACGCAAGTGCAAA' 3'	300
TaCBF11 C21R1L8-a	Forward	5' CACACAAACACCCAACCAACATCAC 3'	320
TaCBF11 C21R1L8-a	Reverse	5' AGCCAGATGCGGGACTTCTTGTT 3'	320
TaCBF12-D L6B007_F24- 19	Forward	5' CCACACTCGTCTACCCAACA 3'	265
TaCBF12-D L6B007_F24- 19	Reverse	5' CTCCTTGAAGTTGGTGCGC 3'	265
TaCBF13-B L6B022_I06- 10	Forward	5' ACACGTCCGGCGAAATGGA 3'	271
TaCBF13-B L6B022_I06- 10	Reverse	5' CCAAATGTCCCGCATTTTATTAT 3'	271
TaCBF14-A BQ744365R1L 8-153	Forward	5' TAGAATTCGAAATGCAGGGCTATC 3'	291
TaCBF14-A BQ744365R1L 8-153	Reverse	5' CACTAATTGTTTGTTGACCCACAT 3'	291

Annexe 3

REFERENCES

- Antikainen, M., and Griffith, M. (1997) Antifreeze protein accumulation in freezing-tolerant cereals. *Physiol. Plant.* 99: 423-432.
- Artus, N.N., Uemura, M., Steponkus, P.L., Gilmour, S.J., Lin, C., and Thomashow, M.F., (1996) Constitutive expression of the cold-regulated *Arabidopsis thaliana* COR15a gene affects both chloroplast and protoplast freezing tolerance. *Proc Natl Acad Sci U S A.* 93: 13404-13409.
- Breton, G., Danyluk, J., Ouellet, F., and Sarhan, F. (2000) Biotechnological applications of plant freezing associated proteins. *Biotechnol Ann Rev.* 6: 57-100.
- Choi, D., Rodriguez, E. M., and Close, T. J. (2002) Barley *cbf3* Genes Identification, Expression pattern, and Map location *Plant Physiol* 129: 1-7.
- Crowe, J.H., Oliver, A.E., Hoekstra, F.A., and Crowe, L.M. (1997) Stabilization of dry membranes by mixtures of hydroxyethyl starch and glucose: the role of vitrification. *Cryobiology.* 35(1): 20-30.
- Crowe, L.M., and Crowe, J.H. (1992) Anhydrobiosis: a strategy for survival. *Adv Space Res.* 12(4): 239-47.
- Danyluk, J., Kane, N. A., Breton, G., Limin, A. E., Fowler, D. B., and Sarhan, F. (2003) TaVRT-1, a putative transcription factor associated with vegetative to reproductive transition in cereals. *Plant Physiol.* 132(4): 1849-60.
- Danyluk, J., Perron, A., Houde, M., Limin, A., Fowler, B., Benhamou, N. and Sarhan, F. 1998. Accumulation of an acidic dehydrin in the vicinity of the plasma membrane during cold acclimation of wheat. *The Plant Cell* 10: 623-638.
- Danyluk, J., Carpentier, E., and Sarhan, F. (1996) Identification and characterization of a low temperature regulated gene encoding an actin-binding protein from wheat. *FEBS Lett.* 389: 324-7.
- Danyluk, J., Houde, M., Rassart, E., and Sarhan, F. (1994) Differential expression of a gene encoding an acidic dehydrin in chilling sensitive and freezing tolerant gramineae species. *FEBS Lett.* 9; 344 (1): 20-4.

Fowler, B., Sarhan, F., Limin, A., and Chauvin, L.P. (1996) The regulatory role of vernalization in the expression of low-temperature induced genes in wheat and rye. *Theor.Appl.Genet.* 93: 554-559.

Gao, M.J., Allard, G., Byass, L., Flanagan, A.M., and Singh, J. (2002) Regulation and characterization of four CBF transcription factors from *Brassica napus*. *Plant Mol Biol* 49: 459-71.

Gilmour, S.J., Sebolt, A.M., Salazar, M.P., Everard, J.D., and Thomashow, M.F. (2000) Overexpression of *Arabidopsis* CBF3 transcriptional activator mimics multiple biochemical changes associated with cold acclimation. *Plant physiol.* 124: 1854-1865.

Gilmour, S.J., Zarke, D.G., Stockinger, E.J., Salazar, M.P., Houghton, J.M., and Thomashow, M.F. (1998) Low temperature regulation of the *Arabidopsis* CBF family of AP2 transcriptional activators as an early step in cold-induced COR gene expression. *Plant J.* 16: 433-442.

Gilmour, S.J., Hajela, R.K., and Thomashow, M.F. (1988) Cold acclimation in *Arabidopsis thaliana*. *Plant Physiol.* 87: 745-750.

Goff, S.A., Ricke, D., Lan, T.H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., Hadley, D., Hutchison, D., Martin, C., Katagiri, F., Lange, B.M., Moughamer, T., Xia, Y., Budworth, P., Zhong, J., Miguel, T., Paszkowski, U., Zhang, S., Colbert, M., Sun, W.L., Chen, L., Cooper, B., Park, S., Wood, T.C., Mao, L., Quail, P., Wing, R., Dean, R., Yu, Y., Zharkikh, A., Shen, R., Sahasrabudhe, S., Thomas, A., Cannings, R., Gutin, A., Pruss, D., Reid, J., Tavtigian, S., Mitchell, J., Eldredge, G., Scholl, T., Miller, R.M., Bhatnagar, S., Adey, N., Rubano, T., Tusneem, N., Robinson, R., Feldhaus, J., Macalima, T., Oliphant, A., and Briggs, S. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica). *Science* 296: 92-100.

Gutterson, N., and Zhang, J.Z., (2004). Genomics applications to biotech traits: a revolution in progress? *Current Opinion in Plant Biology.* 7: 226-230

Gong, Z., Koiwa, H., Cushman, M.A., Ray, A., Bufford, D., Kore-eda, S., Matsumoto, T.K., Zhu, J., Cushman, J.C., Bressan, R.A., and Hasegawa, P.M. (2001) Genes that are uniquely stress regulated in salt overly sensitive (sos) mutants. *Plant Physiol.* 126: 363-75.

Guy, C.L. (1990). Cold acclimation and freezing stress tolerance: role of protein metabolism. *Annu. Rev. Plant. Physiol. Plant. Mol. Biol.* 41: 187-223.

Hong, B., Barg, R., and Ho, T.H. (1992) Developmental and organ-specific expression of an ABA- and stress-induced protein in barley. *Plant Mol Biol.* 18(4): 663-74.

Houde, M., Dhindsa, R.S., and Sarhan, F. (1992) A molecular marker to select for freezing tolerance in Gramineae. *Mol Gen Genet.* 234(1): 43-8.

Ishitani, M., Xiong, L., Stevenson, B., and Zhu J. K. (1997) Genetic analysis of osmotic and cold stress signal transduction in *Arabidopsis*: interactions and convergence of abscisic acid-dependent and abscisic acid-independent pathways. *Plant Cell.* 9 (11):1935-49.

Jaglo, K.R., Kleff, S., Amundsen, K.L., Zhang, X., Haake, V., Zhang, J.Z., Deits, T., and Thomashow, M.F. (2001). Components of the *Arabidopsis* C-repeat/dehydration-responsive element binding factor cold-response pathway are conserved in *Brassica napus* and other plant species. *Plant Physiol.* 127, 910–917.

Jofuku, K.D., den Boer, B.G., Van Montagu, M., and Okamoto, J.K. (1994). Control of *Arabidopsis* flower and seed development by the homeotic gene APETALA2. *Plant Cell.* 6: 1211-1236.

Kagaya, Y., Ohmiya, K., and Hattori, T. (1999). RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. *Nucleic Acids Res.* 27: 470-478.

Lee, S., Huh, K., An, K., An, G., and Kim, S., (2004) Ectopic expression of a Cold-inducible Transcription Factor, CBF1/DREB1b, in Transgenic Rice (*Oryza sativa* L.). *Mol. Cells.* 18: 107-114.

Levitt, J. (1980) Chilling, freezing, and high temperature stress. Academic press, New York. 3-56.

Lin, C., and Thomashow, M.F. (1992) A cold-regulated *Arabidopsis* gene encodes a polypeptide having potent cryoprotective activity. *Biochem Biophys Res Commun.* 183 (3):1103-8.

Lynch, D.V., and Steponkus, P.L. (1987) Plasma membrane lipid alterations associated with cold acclimation of winter rye seedlings (*Secale cereale* L. cv. Puma). *Plant Physiol.* 83: 761-767.

Magnani, E., Sjolander, K., and Hake, S. (2004) From Endonucleases to Transcription Factors: Evolution of the AP2 DNA Binding Domain in plants. *Plant Cell Prev.*

Martin, M.L., and Buscon, L. (2000) Membrane localization of a rice calcium-dependent protein kinase (CDPK) is mediated by myristoylation and palmitoylation. *Plant J.* 2000 Nov;24(4):429-35. Review.

McKersie, B.D., Chen, Y., de Beus, M., Bowley, S.R., Bowler, C., Inze, D., D'Halluin, K., and Botterman, J. (1993) Superoxide dismutase enhances tolerance of freezing stress in transgenic alfalfa (*Medicago sativa* L.). *Plant Physiol.* 103 (4):1155-63.

Medina, J., Bargues, M., Terol, J., Perez-Alonso, M., and Salinas, J. (1999). The Arabidopsis CBF gene family is composed of three genes encoding AP2 domain-containing proteins whose expression is regulated by low temperature but not by abscisic acid or dehydration. *Plant Physiol.* 119: 463-469.

Mizukami, Y., and Fischer, R.L. (2000) Plant organ size control: AINTEGUMENTA regulates growth and cell numbers during organogenesis. *Proc Natl Acad Sci U S A.* 97: 942-7.

Morot-Gaudry, J.F. (2001) Biologie moléculaire et Génomique, Glossaire. Ecole thématique Biologie végétale.

Monroy, A. F., Sarhan, F., and Dhindsa R. S. (1993) Cold-Induced Changes in Freezing Tolerance, Protein Phosphorylation, and Gene Expression (Evidence for a Role of Calcium). *Plant Physiol.* 102(4): 1227-1235.

Okamuro, J.K., Caster, B., Villarroel, R., Van Montagu, M., and Jofuku, K.D., (1997). The AP2 domain of APETALA2 defines a large new family of DNA binding proteins in *Arabidopsis*. *Proc. Natl. Acad. Sci.* 13: 7076-7157.

Ouellet, F., Houde, M., and Sarhan, F. (1993). Purification, characterization and cDNA cloning of the 200 kDa protein induced by cold acclimation in wheat. *Plant Cell Physiol.* 34: 59-65.

Puhakainen, T., Hess, M.W., Makela, P., Svensson, J., Heino, P., and Palva, E.T. (2004) Overexpression of multiple dehydrin genes enhances tolerance to freezing stress in *Arabidopsis*. *Plant Mol Biol.* 54: 743-53.

Qin, F., Sakuma, Y., Li, J., Liu, Q., Li, Y.Q., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2004) Cloning and functional analysis of a novel DREB1/CBF transcription factor involved in cold-responsive gene expression in *Zea mays* L. *Plant Cell Physiol.* 45: 1042-52.

Riechmann, J.L., and Meyerowitz, E.M., (1998). The AP2/EREBP family of plant transcription factors. *J. Bio. Chem.* 379: 633-646.

Sakuma, Y., Liu, Q., Dubouzet, J.G., Abe, H., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2002) DNA-binding specificity of the ERE/AP2 domain of *Arabidopsis* DREBs, transcription factors involved in dehydration and cold-induced gene expression. *Biochem. Biophys. Res. Commun.* 290: 998-1009.

Sarhan, F., Ouellet, F., and Vazquez-Tello, A. (1997) The wheat Wcs120 gene family. A useful model to understand the molecular genetics of freezing tolerance in cereals. *Physiologia Plantarum* 101: 439-445.

Shinozaki, K., and Yamaguchi-shinozaki, K. (2000) Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. *Curr. Opin. Plant Biol.* 3: 217-223.

Shinozaki, K., and Yamaguchi-shinozaki, K. (1996) Molecular responses to drought and cold stress. *Curr Opin Biotechnol.* 7(2): 161-7. Review.

Stockinger, E.J., Mao, Y., Regier, M.K., Triezenberg, S.J., and Thomashow, N.F. (2001) Transcriptional adaptor and histone acetyltransferase proteins in *Arabidopsis* and their interactions with CBF1, a transcriptional activator involved in cold-regulated gene expression. *Nucleic acids research.* 29: 1524-1557.

Stockinger, E.J., Glimour, S.J., and Thomashow, N.F. (1997) *Arabidopsis thaliana* CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit, *Proc. Natl. Acad. Sci. USA* 94: 1035-1040.

Tao, D., Öquist, G., and Wingsle, G. (1998) Active Oxygen Scavengers during Cold Acclimation of Scots Pine Seedlings in Relation to Freezing Tolerance. *Cryobiology.* 37(1): 38-45.

Thomashow, M.F. (1999) Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. *Ann Rv Plant Physiol Plant. Mol. Biol.* 50: 571-599.

Thomashow, M.F. (1998). Role of cold-responsive genes in plant freezing tolerance. *Plant Physiol.* 118: 1-7.

Thomashow, M. F. (1990). Molecular genetics of cold acclimation in higher plants. *Adv Genet.* 28: 99-131.

Tournier, B., Sanchez-Ballesta, M.T., Jones, B., Pesquet, E., Regad, F., Latche, A., Pech, J.C., and Bouzayen, M. (2003) New members of the tomato ERF family show specific expression patterns and diverse DNA-binding capacity to the GCC box element. *FEBS Lett.* 550: 149-54.

Uemura, M., Joseph, R.A., and Steponkus, P.L. (1995) Cold acclimation of *Arabidopsis thaliana* (Effect on plasma membrane lipid composition and freeze-induced lesions). *Plant Physiol.* 109: 15-30.

Uemura, M., and Steponkus, P.L. (1994) A Contrast of the Plasma Membrane Lipid Composition of Oat and Rye Leaves in Relation to Freezing Tolerance. *Plant Physiol.* 104(2): 479-496.

Vahala, T., Oxelman, B., and von Arnold, S. (2001) Two APETALA2-like genes of *Picea abies* are differentially expressed during development. *J Exp. Bot.* 52: 1111-1116.

van de Peer Y, de Wachter R (1994) TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Comput Appl Biosci* 10: 569-570.

Xiong Y. and Fei S-Z. (2006) Functional and phylogenetic analysis of a DREB/CBF-like gene in perennial ryegrass (*Lolium perenne* L.). *Planta*

Xue, G.P. (2003) The DNA-binding activity of an AP2 transcriptional activator HvCBF2 involved in regulation of low-temperature responsive genes in barley is modulated by temperature. *Plant J.* 33 (2): 373-83.

Yamaguchi-Shinozaki, K., and Shinozaki, K. (1997) Gene Expression and Signal Transduction in Water-Stress Response. *Plant Physiol.* 115 (2): 327-334.

Zhang, X., Fowler, S.G., Cheng, H., Lou, Y., Rhee, S.Y., Stockinger, E.J., and Thomashow, M.F. (2004) Freezing-sensitive tomato has a functional CBF cold

response pathway, but a CBF regulon that differs from that of freezing-tolerant *Arabidopsis*. *Plant J.* 39: 905-924.